



UNIVERSIDAD DE MÁLAGA



Ingeniería de la Salud

Uso de técnicas de Visión por Computador y Aprendizaje Profundo en dosimetría biológica para la detección de cromosomas dicéntricos en placas metafásicas

Use of Computer Vision and Deep Learning techniques in biological dosimetry for dicentric chromosome detection in metaphasic plates

Realizado por  
Ignacio Atencia Jiménez

Tutorizado por  
Miguel Ángel Molina Cabello  
Miguel José Ruíz Gómez

Departamento  
LENGUAJES Y CIENCIAS DE LA COMUNICACIÓN  
UNIVERSIDAD DE MÁLAGA

MÁLAGA, Febrero 2024



ESCUELA TÉCNICA SUPERIOR DE INGENIERÍA  
INFORMÁTICA  
GRADO EN INGENIERÍA DE LA SALUD

**Uso de técnicas de Visión por Computador y Aprendizaje  
Profundo en dosimetría biológica para la detección de  
cromosomas dicéntricos en placas metafásicas**

**Use of Computer Vision and Deep Learning  
techniques in biological dosimetry for dicentric  
chromosome detection in metaphasic plates**

Realizado por  
**Ignacio Atencia Jiménez**

Tutorizado por  
**Miguel Ángel Molina Cabello**  
**Miguel José Ruíz Gómez**

Departamento  
**Lenguajes y Ciencias de la Comunicación**

UNIVERSIDAD DE MÁLAGA  
MÁLAGA, FEBRERO DE 2024

Fecha defensa: Marzo de 2024

*A mis padres, M<sup>a</sup>José y José, por facilitarme la formación que ellos no pudieron tener, y por haberme acompañado en todas las etapas durante este emocionante proceso. Y a mi abuela Fina, por siempre escuchar con interés cada avance en este Trabajo Fin de Grado (aun a veces sin entender del todo sobre lo que le estaba contando).*

# Abstract

Dicentric chromosome counting is a technique widely used in biological dosimetry to quantify the genetic damage caused by radiation exposure and to estimate the dose the subject has received. Currently, the technique is performed manually. However, interest has recently focused on automating this process using Artificial Intelligence tools to reduce analysis time and improve the quality of the results. This project aims to contribute to the state of the art of this task for which, to date, there is still little research.

In this work, a convolutional neural network has been trained from a set of images (labelled by hand) on which different preprocessing techniques were applied. This model has been trained for each of the experiments, which represents a detailed study of the problem that had never been carried out before.

The predictions obtained for each of the experiments have shown a good performance in chromosome detection. Likewise, the majority of non-dicentric chromosomes are recognised. However, the model presents difficulties in characterising dicentric chromosomes.

**Keywords:** Deep Learning, Computer Vision, Biological Dosimetry, Dicentric Chromosomes, Radiation

# Resumen

El recuento de cromosomas dicéntricos es una técnica ampliamente utilizada en dosimetría biológica para cuantificar el daño genético producido por la exposición a radiación y estimar la dosis que el sujeto ha recibido. Actualmente, la técnica se realiza de forma manual. No obstante, recientemente el interés se ha centrado en lograr la automatización de dicho proceso a través de herramientas de Inteligencia Artificial para lograr una reducción del tiempo de análisis y una mejora en la calidad de los resultados. Este proyecto pretende contribuir al estado del arte de esta tarea de la que, a día de hoy, aún existe poca investigación.

En este trabajo se ha entrenado una red neuronal convolucional a partir de un conjunto de imágenes (etiquetadas a mano) sobre las que se aplicaron diferentes técnicas de preprocesado. Dicho modelo ha sido entrenado para cada uno de los experimentos, lo que supone un estudio pormenorizado del problema que nunca antes se había llevado a cabo.

Las predicciones obtenidas para cada uno de los experimentos han mostrado un buen desempeño en la detección de los cromosomas. Asimismo, los cromosomas no dicéntricos son reconocidos en su mayoría. No obstante, el modelo presenta dificultades caracterizando a los cromosomas dicéntricos.

**Palabras clave:** Aprendizaje Profundo, Visión por Computador, Dosimetría Biológica, Cromosomas Dicéntricos, Radiación



# Índice

<b>1. Introducción</b>	<b>7</b>
1.1. Justificación . . . . .	9
1.2. Objetivos . . . . .	9
1.3. Metodología de Trabajo . . . . .	10
1.4. Tecnologías empleadas . . . . .	10
1.5. Estructura del documento . . . . .	12
<b>2. Fundamentos Biológicos</b>	<b>13</b>
2.1. Ácido Desoxirribonucleico (DNA) . . . . .	13
2.2. Ciclo Celular: Interfase y mitosis . . . . .	16
2.3. Efectos de la radiactividad y dosimetría biológica . . . . .	18
<b>3. Contexto técnico</b>	<b>25</b>
3.1. Estado del Arte . . . . .	25
3.1.1. Aprendizaje Profundo . . . . .	25
3.1.2. Visión por Computador . . . . .	26
3.2. Proceso de Aprendizaje . . . . .	27
3.2.1. Pasada hacia adelante . . . . .	28
3.2.2. Retropropagación . . . . .	30
3.2.3. Red neuronal convolucional . . . . .	33
3.2.4. Hiperparámetros . . . . .	35
3.3. Proceso de Visión . . . . .	36
<b>4. Resultados y Discusión</b>	<b>39</b>
4.1. Métricas . . . . .	39
4.2. Preprocesado de Imágenes . . . . .	42
4.2.1. Técnicas de Umbralización . . . . .	42
4.2.2. Filtros morfológicos . . . . .	45
4.3. Etiquetado de imágenes . . . . .	47

4.4. Validación Cruzada . . . . .	48
4.5. Entrenamiento . . . . .	49
4.6. Ensamblado . . . . .	52
4.7. Discusión . . . . .	56
<b>5. Conclusiones y Líneas Futuras</b>	<b>59</b>
5.1. Conclusiones . . . . .	59
5.2. Líneas Futuras . . . . .	60

# 1

## Introducción

La dosimetría biológica es una disciplina que se encarga de la medición indirecta de la radiación ionizante absorbida por un individuo. Una de sus técnicas consiste en el conteo de cromosomas dicéntricos en placas metafásicas de linfocitos. Los cromosomas dicéntricos son cromosomas que, debido a los efectos de la radiación ionizante, han experimentado roturas en los brazos de sus cromátidas y que, en el proceso de reparación, han quedado fusionados con otros cromosomas afectados, dando lugar a cromosomas con dos centrómeros.

La frecuencia de dicéntricos permite establecer una relación con la radiación absorbida por un individuo que haya sufrido una exposición. A partir de estos valores, el personal especializado puede determinar si el paciente requiere o no de tratamiento curativo, o, en el peor de los casos, si requiere de tratamiento paliativo.

Actualmente, es el propio personal sanitario quien sirve de operador para determinar la frecuencia de dicéntricos a partir de las imágenes obtenidas en placas metafásicas. Esto tiene dos claras implicaciones: baja ineficiencia en la obtención de resultados (debido a la elevada cantidad de tiempo que requiere el análisis) y resultados sometidos al sesgo del operador. Ante esta situación, se hace patente la necesidad de automatizar dicho proceso mediante técnicas de inteligencia artificial.

La inteligencia artificial consiste en la capacidad de las máquinas de resolver problemas a través de algoritmos que imitan a los procesos que subyacen en la corteza cerebral humana y que dan lugar a aquello que se identifica por inteligencia humana. De entre estos procesos, caben destacar dos: la visión y el aprendizaje.

La Visión por Computador (en inglés, *Computer Vision*) es una rama de la inteligencia artificial que se dedica a la detección espacial de objetos en imágenes y videos. Por su parte, el Aprendizaje Profundo (en inglés, *Deep Learning*) se centra en la abstracción de patrones de un conjunto de datos de entrada con el fin de poder hacer predicciones sobre datos distintos a

estos.

Estas técnicas se basan en el uso de redes neuronales convolucionales, un tipo de arquitectura computacional que consiste en la aplicación sucesiva de filtros sobre una serie de capas de neuronas conectadas entre sí. Cada filtro aplica una corrección matemática a la imagen que permite aumentar el grado de abstracción conforme se avanza a través de las capas de neuronas. La corrección aplicada por cada filtro viene determinada por sus pesos y por los sesgos de las neuronas de salida, los cuales se modifican a lo largo del proceso de aprendizaje a fin de conseguir predicciones óptimas. La salida de la red corresponde con la predicción para la imagen.

Para poder desempeñar correctamente su labor en la detección de un determinado objeto en una imagen, las redes neuronales convolucionales deben ser previamente entrenadas con un conjunto de imágenes etiquetadas. Una etiqueta es un documento que registra cada uno de los objetos de una imagen que se espera que el modelo sea capaz de identificar. En dicho documento, se deben especificar tanto la clase del objeto como las coordenadas  $x$  e  $y$  que definen al mínimo recuadro que encierra al objeto (en inglés, *bounding box*).

El entrenamiento consiste en una progresiva modificación de los pesos y sesgos de cada neurona en múltiples etapas con el fin de minimizar el error cometido en cada predicción. En consecuencia, se consiguen salidas cada vez más acertadas.

El proceso de validación pone a prueba al modelo en cada una de las etapas del entrenamiento con imágenes ajenas a las utilizadas en este último proceso. Permite llevar un registro de la evolución del desempeño del modelo en las diferentes etapas. Son las métricas arrojadas por el proceso de validación las que permiten discriminar entre modelos mejores y peores para una determinada tarea.

El proceso de pruebas permite hacer una última criba en la que se introduce una nueva imagen en el modelo ya previamente entrenado y validado. El modelo devuelve la predicción sobre dicha imagen y visualmente se puede comprobar si, tras el entrenamiento, el modelo es capaz de reconocer objetos en imágenes con acierto.

Por otra parte, existe una serie de características de la red, los hiperparámetros, que pueden ser establecidos previamente por el programador. Determinan el funcionamiento y estructura de la red y, consecuentemente, condicionan su desempeño. El estudio del funcionamiento del modelo ante distintas configuraciones de hiperparámetros permite el descarte de modelos con

peor operatividad, así como la selección de modelos más refinados.

## 1.1. Justificación

El Consejo de Seguridad Nuclear (CSN) es un organismo estatal cuyas acciones se dirigen al estudio de la protección radiológica y la seguridad nuclear con tal de prever y corregir cualquier tipo de emergencia radiológica. Una de las líneas prioritarias de financiación del CSN en 2023 es la automatización de las técnicas de dosimetría biológica [1].

Actualmente, el recuento de cromosomas dicéntricos en placas metafásicas requiere de un profesional de la Salud, que debe invertir una determinada cantidad de tiempo en localizar los cromosomas anómalos. Esto puede llegar a convertirse en una tarea ardua si se considera que de cada paciente se han de evaluar múltiples imágenes de placas metafásicas para poder extraer conclusiones aceptables.

Es por ello que la automatización de este proceso permitiría cuantificar la dosis absorbida por los trabajadores expuestos a radiación ionizante de una forma mucho más eficiente. Asimismo, podría contribuir a evitar un colapso de los hospitales en caso de un hipotético accidente radiológico que provoque la exposición de un gran grupo de la población.

## 1.2. Objetivos

El principal objetivo de este trabajo es la detección de cromosomas dicéntricos en placas metafásicas mediante el desarrollo de *software* de inteligencia artificial basado en Aprendizaje Profundo y Visión por Computador. A tal fin, se elaborará código en lenguaje de programación Python que permita el entrenamiento de un modelo de predicción. El modelo, en último término, habrá de ser capaz de detectar cada uno de los cromosomas de la imagen que se pase como entrada y discernir entre cromosoma alterado y cromosoma no alterado.

Para lograr el objetivo global de este trabajo se plantean los siguientes objetivos específicos:

- Estudio de los fundamentos biológicos. Se hará una búsqueda bibliográfica que permita contextualizar las bases biológicas sobre las que se asienta este estudio.
- Estudio del contexto técnico. Se hará una búsqueda bibliográfica sobre Visión por Computador y Aprendizaje Profundo.

- Elaboración del código. Se preprocesarán las imágenes y se entrenará el modelo con diferentes experimentos.
- Discusión de los resultados. Se comentarán los resultados obtenidos.

### 1.3. Metodología de Trabajo

Los objetivos previamente desarrollados se alcanzarán a través de la realización de código aplicando la metodología ágil y el método científico. La metodología ágil es un método iterativo caracterizado por adaptar las etapas de trabajo a las condiciones concretas del proyecto en el que se trabaja, donde la comunicación entre las partes implicadas y la retroalimentación son un elemento indispensable. Por su parte, el método científico es una metodología basada en la creación de nuevo conocimiento a partir de la experimentación y la formulación de hipótesis, donde la refutabilidad, la reproducibilidad y la repetibilidad son elementos indispensables.

### 1.4. Tecnologías empleadas

Para la realización de este trabajo se han de considerar una serie de herramientas *software* que permitan el trabajo con inteligencia artificial, así como un editor de textos que permita la elaboración de la memoria. A continuación, se procede a enumerar el conjunto de tecnologías a emplear:

- Phyton: Lenguaje de programación de alto nivel con multitud de recursos para el trabajo con inteligencia artificial.
- sys: Librería que permite la manipulación de los parámetros de ejecución.
- zipfile: Librería que permite el manejo de archivos .zip.
- NumPy: Librería que proporciona herramientas para la operación con vectores y matrices.
- Pandas: Librería que proporciona herramientas para la manipulación y análisis de datos.
- os : Librería que proporciona herramientas para el manejo de directorios.

- shutil : Librería que proporciona herramientas para la operación con archivos y directorios en alto nivel.
- Sci-Py: Librería que proporciona herramientas para la manipulación de señales.
- Sci-Kit Learn: Librería que proporciona herramientas de Aprendizaje Computacional.
- Matplotlib: Librería para la generación de gráficos.
- OpenCV: Librería que proporciona herramientas para el preprocesado de imágenes.
- YOLO: Paquete que permite el entrenamiento de modelos aplicando Aprendizaje Profundo y Visión por Computador.
- PhyCharm: Entorno de desarrollo local para Python.
- Google Drive: Servicio de alojamiento de archivos en la nube.
- Google Colab: Entorno de desarrollo online para Python con posibilidad de uso de GPUs.
- CVAT: Herramienta online para el etiquetado de imágenes.
- Adobe Photoshop: Editor de imágenes.
- GitHub: Plataforma de desarrollo colaborativo que permite el alojamiento de proyectos en la nube.
- Git: Sistema de control de versiones empleado por GitHub.
- LaTeX: Sistema de composición de textos.
- Overleaf: Editor de textos online.

Asimismo, también cabe destacar que para el estudio se ha empleado un conjunto de 50 micrografías de cromosoma privadas cedidas por el centro Radiation Emergency Assistance / Training Site (REAC/TS) del Instituto *Oak Ridge Institute for Science and Education* (ORISE) de Tennessee, Estados Unidos y por la Dra. Alegría Montoro Pastor del Hospital la Fe de Valencia. Además, se ha utilizado la arquitectura yolov8x como modelo de entrenamiento.

## 1.5. Estructura del documento

La presente sección permitirá al lector ubicar los distintos contenidos que se desarrollan a lo largo de esta memoria. Los capítulos y los aspectos tratados en cada uno de ellos son los que siguen:

**CAPÍTULO 2, FUNDAMENTOS BIOLÓGICOS:** Una breve introducción a los fundamentos en dosimetría biológica que permitirán definir el ámbito y contexto del problema que se pretende resolver. **CAPÍTULO 3, CONTEXTO TÉCNICO:** Un repaso a los principales conceptos clave e hitos en inteligencia artificial que servirán de punto de partida para abordar el problema. **CAPÍTULO 4, RESULTADOS Y DISCUSIÓN:** Exposición de los resultados del trabajo y discusión sobre aspectos relacionados con el mismo. **CAPÍTULO 5, CONCLUSIONES Y LÍNEAS FUTURAS:** Breve exposición de las conclusiones finales, así como un vistazo hacia las posibles actuaciones que permitan la evolución de este trabajo.

# 2

# Fundamentos Biológicos

## 2.1. Ácido Desoxirribonucleico (DNA)

El ácido desoxirribonucleico, o DNA por sus siglas en inglés (*desoxirribonucleic acid*), es una molécula descrita por primera vez por Watson y Crick en la revista Nature en 1953 [2]. Pertenece a la categoría de ácidos nucleicos, uno de los cuatro grupos de biomoléculas más estudiados.

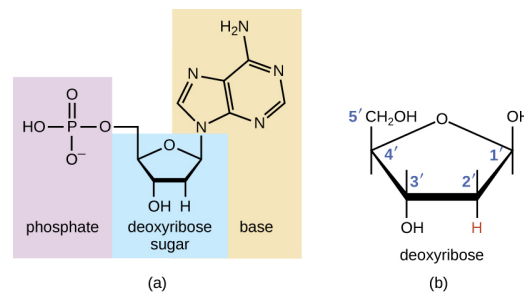


Figura 1: Componentes elementales del DNA [3]. (a) Estructura de un desoxirribonucleótido. (b) Desoxirribosa

Su estructura consiste en cadenas de desoxirribonucleótidos (véase Figura 1a). Los desoxirribonucleótidos presentan en su estructura una molécula glucídica de cinco carbonos: la desoxirribosa (véase Figura 1b).

Además, la desoxirribosa se caracteriza por carecer de un grupo funcional hidroxilo en su segundo carbono (a diferencia de la molécula de ribosa, que sí presenta grupo hidroxilo en su segundo carbono y que da lugar a cadenas de ácido ribonucleico, o RNA, por sus siglas en inglés) y se une por su primer carbono a una base nitrogenada a través de un enlace N-Glucosídico. Dicha base nitrogenada puede ser una purina (guanina o adenina) o una piri-

midina (citosina o timina).

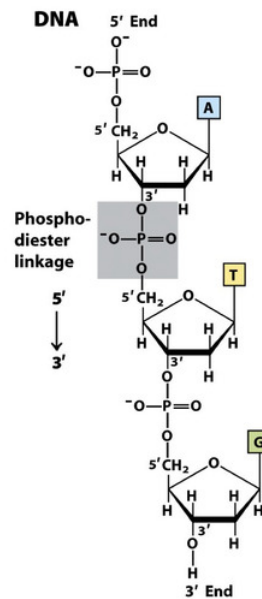


Figura 2: Estructura primaria del DNA [4]. Nótese el enlace nucleotídico (gris) y la polaridad de la cadena (extremo 3' y extremo 5')

En su quinto carbono, la desoxirribosa presenta un grupo fosfato que le permite unirse a otro nucleótido a través del tercer carbono de este último mediante un enlace fosfodiéster conocido como enlace nucleotídico (véase la Figura 2). Esta particularidad en la unión denota cierta polaridad en las cadenas de DNA, permitiendo así discernir cada uno de los extremos de la cadena como extremo 3' (el nucleótido terminal se une a otro nucleótido por su quinto carbono, quedando el tercer carbono libre) o extremo 5' (el nucleótido terminal se une a otro nucleótido por su tercer carbono, quedando el quinto carbono libre). El resto de nucleótidos que no se encuentran en los extremos de la cadena presentan uniones tanto en el tercer carbono como en el quinto.

Por otra parte, dos cadenas de nucleótidos pueden unirse para dar lugar a la molécula típica de DNA tan presente en el imaginario colectivo, la que se puede apreciar en la Figura 3. Para que tenga lugar esta unión, se deben dar dos condiciones: las cadenas han de ser complementarias y antiparalelas ente sí. Esta situación permite la unión de bases nitrogenadas complementarias a través de dos puentes de hidrógeno en el caso de la unión adenina-timina (A-T) y tres puentes de hidrógeno en el caso de la unión guanina-citosina (C-G).

La secuencia de nucleótidos, o más concretamente, la combinación de bases nitrogenadas,

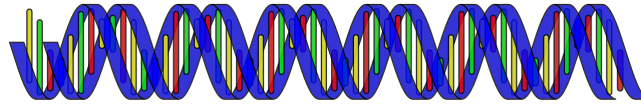


Figura 3: Molécula de DNA [5]. En azul: cadena de nucleótidos. En amarillo, verde y rojo: uniones entre bases nitrogenadas (A-T, C-G)

codifican un determinado mensaje que puede ser transcrito a RNA para posteriormente ser traducido a proteína. A cada una de las secuencias que codifican para una proteína se las conoce como gen. Es por tanto, el DNA, la molécula que presenta la información necesaria para la producción de todas las proteínas del organismo.

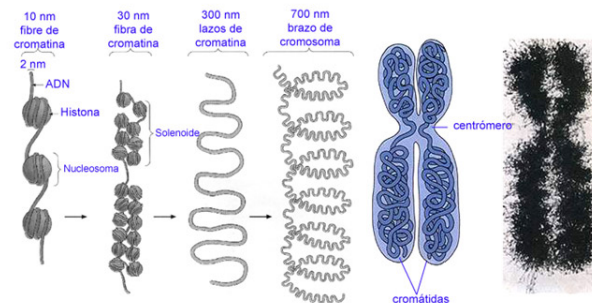


Figura 4: Grados de compactación del DNA [6].

El DNA se encuentra en el interior del núcleo de las distintas células del organismo, así como en el interior de las mitocondrias (aunque este trabajo se centra en el DNA nuclear). No suele presentar la estructura secundaria de doble cadena previamente descrita, sino que suele aparecer con un mayor grado de compactación (véase la Figura 4).

Para poder encerrarse en el núcleo de una célula, la extensa doble cadena de DNA debe enrollarse en torno a una serie de proteínas estructurales conocidas como histonas, dando lugar a estructuras con mayor empacamiento: la cromatina.

En la célula interfásica, existen determinadas regiones donde dicha cromatina puede encontrarse extendida y otras regiones donde aparece condensada. Las regiones más compactas protegen a determinados genes de ser transcritos por RNA-Polimerasas, esto es, los previenen de ser expresados. Este mecanismo resulta de gran relevancia para evitar que ciertas proteínas sean sintetizadas por un tipo celular al que este proceso no les corresponde. Es importante en este punto recalcar que la información contenida en el DNA es exactamente idéntica en todos

los tipos celulares, pero difiere el grado de empacamiento en el DNA de unos u otros genes (en otras palabras, cuáles genes se expresan y cuáles no).

Un aún mayor grado de compactación es también imprescindible cuando la célula entra en estado de división a fin de proteger el material genético, tal y como se detallará en el siguiente epígrafe.

## 2.2. Ciclo Celular: Interfase y mitosis

Se conoce como ciclo celular a un conjunto de fases que atraviesa la célula desde que es originada a partir de una división celular hasta que esta se vuelve a dividir. Resulta interesante mencionar que, ontogenéticamente, la célula primera es el cigoto, mientras que un enfoque filogenético requeriría retroceder algunos millones de años atrás en el tiempo (aún se teoriza sobre cuál puede ser el origen de la primera célula).

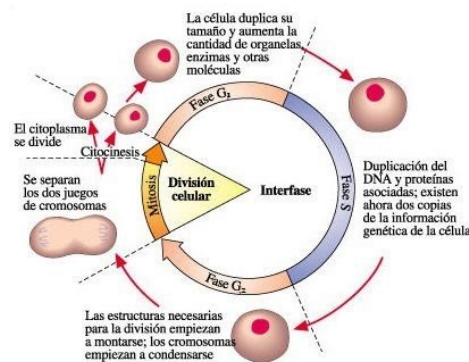


Figura 5: Fases del ciclo celular. [7]

A grandes rasgos, se pueden diferenciar dos fases del ciclo celular: interfase, o fase de crecimiento, y mitosis, o fase de división (véase la Figura 5). Si bien existe otro tipo de división celular, la meiosis (que se da en las células germinales), no se tendrá en consideración puesto que este estudio se centra en imágenes de placas metafásicas de linfocitos (células somáticas).

Durante la interfase, la célula crece en tamaño mientras desarrolla sus funciones. Dado el momento, la célula entra en fase de división a través de un proceso mediado por comunicación celular (activación de los complejos ciclina-CDK).

La replicación del DNA es una etapa en torno al final de la interfase en la que la célula crea una copia de su DNA. Esto es posible gracias a la acción de enzimas DNA-Polimerasas que

sintetizan una nueva cadena de DNA a partir de la adición de nucleótidos complementarios a la cadena que se está duplicando (cadena molde).

El cromosoma es el último estadio de compactación del DNA. Supone un empacamiento de la cromatina mediado por unas proteínas denominadas condensinas, las cuales son activadas por ciclinas M al final de la interfase. Cada una de las copias de DNA se compacta de cromatina a cromosomas. En este punto, cada uno de los cromosomas tiene una copia exacta de sí mismo, producto de la replicación. A cada una de las copias se le conoce como cromátida hermana.

Las cromátidas hermanas se unen entre sí mediante proteínas denominadas cohesinas, las cuales son activadas por ciclinas M al final de la interfase. Dichas cohesinas unen a las cromátidas hermanas en un punto, denominado centrómero (véase de nuevo la Figura 4), dando lugar así a un total de veintitrés pares de cromosomas en una célula en proceso de división.

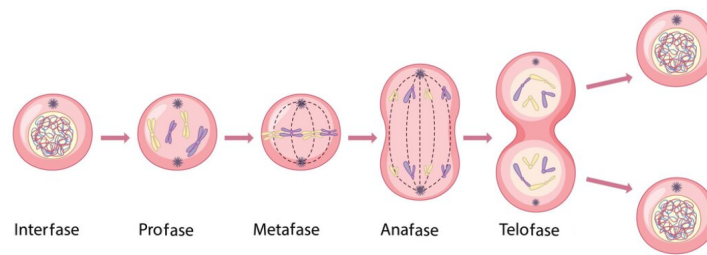


Figura 6: Fases de la mitosis [8]. La citocinesis ocurre tras la telofase, mientras que la prometafase ocurre entre la profase y la metafase.

La mitosis (ilustrada en la Figura 6) consiste en la división del núcleo celular asegurando un idéntico reparto de la información genética en las células hijas (cariocinesis) y posterior división del citoplasma (citocinesis). Las fases de la cariocinesis se pueden resumir como sigue:

- Profase: Compactación de la cromatina en cromosomas mediante condensinas y cohesinas previamente activadas por ciclinas M.
- Prometafase: Desintegración de la membrana nuclear.
- Metafase: Disposición de cromosomas en torno al huso mitótico.
- Anafase: Separación de las cromátidas hermanas, cada una de las cuales acabará en un extremo distinto del huso mitótico.

- Telofase: Formación de membrana nuclear en torno a las cromátidas.

Este estudio utiliza imágenes de placas metafásicas debido a que la disposición de los cromosomas en torno al uso mitótico permite distinguirlos claramente.

En otro orden de cosas, es importante mencionar que el ciclo celular es un proceso altamente regulado con el fin de evitar la transmisión de errores. A este respecto, existen numerosos puntos de control en los que el ciclo celular se detendrá en caso de que las condiciones no sean favorables (ausencia de nutrientes, falta de factores de crecimiento, condiciones ambientales adversas, etc), pudiendo entrar la célula en un estado de quiescencia (detención reversible del ciclo celular, mantenida hasta la aparición de condiciones favorables).

Tanto la entrada a replicación como la entrada a mitosis suponen puntos críticos. En caso de que el material genético esté altamente dañado llegado a dichos puntos críticos, las concentraciones de las moléculas ATM y ATR aumentarán. Esto activará la proteína p53, que desencadenará la vía intrínseca apoptótica, cuyo último término es el suicidio celular (apoptosis) [9].

### **2.3. Efectos de la radiactividad y dosimetría biológica**

La radiactividad es un proceso de emisión de energía en forma de partículas u ondas electromagnéticas por parte de un átomo que se da en átomos inestables a fin de alcanzar un estado estable. Cuanto mayor sea la energía de la radiación, mayor será la capacidad de alterar las moléculas con las que entra en contacto. Concretamente, aquellas radiaciones capaces de producir roturas de enlaces químicos e ionizar los átomos, se denominan radiaciones ionizantes.

Por otra parte, se definen dos unidades de medida para la cuantificación de la dosis absorbida de radiación: Gray (Gy) y Sievert (Sv). Los Grays corresponden con la cantidad de radiación ionizante absorbida sin ponderar por el tipo de radiación u órgano al que afecta. Por su parte, los Sieverts, corresponden con la dosis equivalente, esto es, la radiación absorbida considerando el tipo de radiación y la dosis efectiva cuando se pondera en función de la sensibilidad del tejido irradiado.

Se ha mencionado que la radiación puede presentarse en forma de ondas electromagnéticas o en forma de partículas (véase la Figura 7). A esta última se la conoce por radiación corpuscular. Dicha radiación corpuscular presenta mayor poder ionizante, pero debido a su elevada

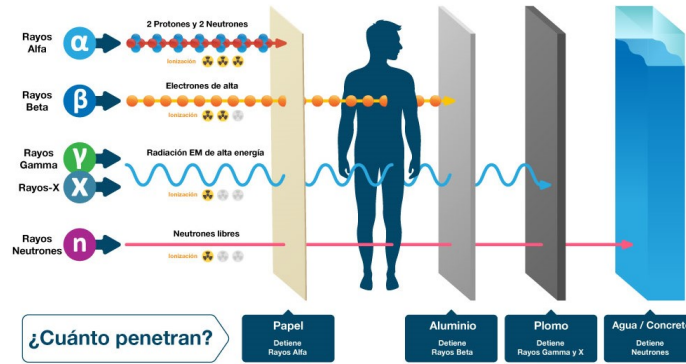


Figura 7: Comparación del poder ionizante y penetración de distintos tipos de radiación, así como su atenuación por distintos materiales [10].

masa, tiene menor grado de penetración. También, se distinguen dos tipos de radiaciones corpusculares: partículas  $\alpha$  (emisión de núcleos de helio cargados positivamente) y partículas  $\beta$  (emisión de electrones resultado de la desintegración de neutrones).

Las radiaciones no corpusculares, al emitirse en forma de ondas, presentan mayor capacidad de penetración. Caben destacar dos: rayos X (emisión de ondas por el frenado de electrones de la capa externa del átomo) y rayos  $\gamma$  (emisión de ondas por desintegración de núcleos inestables).

La alta energía contenida en la radiación ha resultado de gran utilidad para el ser humano, quien ha aprendido a utilizarla en aplicaciones múltiples, tales como la adquisición de imágenes médicas (como son los casos de la radiografía, la radioscopia o la medicina nuclear), tratamientos contra el cáncer o el aprovechamiento de la energía derivada de la fisión de núcleos (energía nuclear) para la obtención de otros tipos de energía: eléctrica, térmica, mecánica... No obstante, errores en la manipulación de núcleos radiactivos, así como otros usos menos ortodoxos, han provocado que cierta parte de la población haya estado expuesta a altas dosis de radiación a lo largo de la Historia.

La radiación ionizante altera la carga de las moléculas, provocando importantes cambios estructurales (que afectan a su función) o incluso llegando a romperlas. Asimismo, la radiación provoca la radiólisis del agua al entrar en contacto con la misma, generando especies reactivas de oxígeno (aniones hidróxido). Si bien la radiación no ionizante no presenta energía como para destruir moléculas, múltiples trabajos han estudiado efectos aditivos a los daños en el

DNA causados por otros agentes mutagénicos [11] , así como envejecimiento acelerado en células de *Saccharomyces cerevisiae* [12].

Por otra parte, el DNA es una molécula altamente sensible a radiación. Como se mencionó anteriormente, esta molécula contiene información para codificar todas las proteínas del organismo, por lo que alteraciones en dicha molécula pueden resultar fatales. Tanto la radiación per se como los radicales libres generados por la misma ocasionan efectos adversos en el DNA. Cabe mencionar dos tipos de daño [13]:

- Rotura de la cadena de DNA: Consiste en la ruptura del enlace nucleotídico entre dos nucleótidos debido a la exposición de la molécula a radiación ionizante, que puede darse en una única hebra o en ambas. Una ruptura en ambas hebras da lugar a la fragmentación de la molécula, siendo esta alteración la causa generadora de cromosomas dicéntricos.
- Alteración de las bases nitrogenadas: Ocurre de forma directa por radiación o por la interacción de especies reactivas del oxígeno con las bases nitrogenadas. Pueden resultar en la escisión, modificación o dimerización de bases nitrogenadas. Se cambia el contenido del mensaje contenido en el DNA.

La maquinaria celular presenta múltiples mecanismos para contrarrestar dichos daños en el DNA a fin de garantizar una correcta transmisión de la información genética a las células hijas.

En algunos casos, los daños en la doble hebra de DNA pueden acabar reflejándose a nivel cromosómico. Fragmentos de cromosomas pueden acabar desprendidos como resultado de múltiples roturas de doble cadena. Como previamente se comentó, la maquinaria celular trata de compensar estas anomalías.

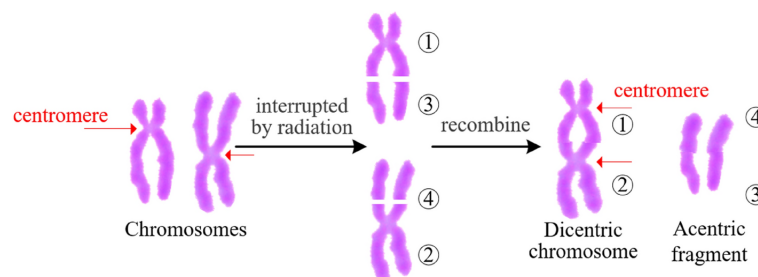


Figura 8: Proceso de formación de cromosomas dicéntricos [14]

Si dos cromosomas se fragmentan producto de la radiación ionizante en algún momento previo a la replicación, lo ideal es que los mecanismos de reparación fusionen cada fragmento con su cromosoma fragmentado. No obstante, si ambos cromosomas están próximos entre sí, también puede producirse un fallo en el proceso de reparación en el que los fragmentos no se fusionen con su cromosoma fragmentado, sino que los cromosomas fragmentados se unen entre sí (véase la Figura 8). Asimismo, los fragmentos también se unen entre sí. Resultan de este modo dos tipos de cromosomas: los derivados de la unión de los cromosomas fragmentados, con dos centrómeros (cromosomas dicéntricos) y los derivados de la unión de fragmentos, sin centrómero (cromosomas acéntricos). En la Figura 9 se muestra un ejemplo del tipo de micrografía con las que se trabajará en este estudio.

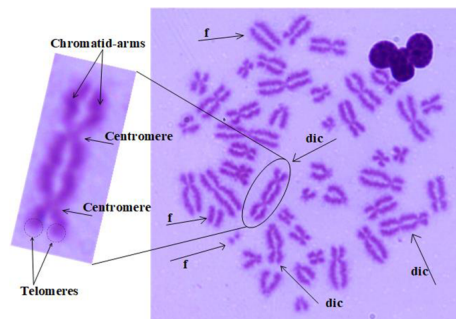


Figura 9: Cromosoma dicéntrico en micrografía de placa metafásica. Se indican con flechas: cromosomas dicéntricos (dic) y cromosomas acéntricos (f). [15]

En otro orden de cosas, los daños que la radiación ejerce sobre el material genético –cuando la exposición supera a la capacidad de los mecanismos de reparación genética– tienen claros efectos biológicos en las células.

La mutación de determinados genes altera proteínas involucradas en la regulación del ciclo celular. Entre estos genes se incluyen los oncogenes (genes que codifican para proteínas que inducen el ciclo celular) y genes supresores de tumores (genes que codifican para proteínas que inhiben el ciclo celular). Mientras que las mutaciones en oncogenes suelen derivar en la expresión excesiva de sus productos, las mutaciones en los genes supresores de tumores suelen derivar en la expresión deficiente –o nula– de sus productos.

En suma, la alteración de oncogenes y genes supresores de tumores afectan al ciclo celular, pudiendo llegar a originar células que se dividen de forma descontrolada, esto es, células cancerígenas.

Tabla 1: Resumen de las dosis y causas de algunos efectos no estocásticos [16].

Tejido	Efecto	Umbral (Gy)	Causa
Sistema hematopoyético	Hemorragias	0.5	Plaquetopenia
Sistema inmune	Inmunodepresión	0.1	Linfopenia
Sistema Gastrointestinal	Infección sistémica	2.0	Lesión del epitelio intestinal
Piel	Descamación	3.0	Daño en la capa basal
Testículo	Esterilidad	0.2	Aspermia celular

La aparición de células cancerígenas no se puede determinar a partir de la cantidad de radiación absorbida, ya que depende en gran medida de qué genes son afectados y el nivel de daño. Más bien se puede estimar con qué probabilidad pueden acaecer. Es por ello que se los conoce por efectos estocásticos.

No obstante, alcanzado cierto umbral de radiación se da en cualquier caso la muerte celular por catástrofe mitótica. El DNA está tan dañado que los puntos de control fallan en detener el ciclo celular y la célula entra en mitosis. A fin de evitar las consecuencias de transmitir material genético aberrante a las células descendientes, la célula entra o bien en estado de senescencia (cese del ciclo celular y funcional permanente e irreversible) o bien de suicidio celular (apoptosis). Cuando la radiación provoca muerte celular en los tejidos, aparecen daños tras un umbral de dosis bien determinado. Estos daños no probabilísticos se denominan daños deterministas o no estocásticos. Algunos efectos no estocásticos de la radiación se ilustran en la Tabla 1.

Los efectos biológicos no tienen por qué aparecer de forma instantánea. Es por ello que detectar a tiempo si un individuo ha recibido radiación puede resultar decisivo en su supervivencia (en tanto que se le pueda suministrar un tratamiento que palie los efectos).

De esta forma surge la dosimetría biológica, que recoge un conjunto de técnicas que permiten determinar si un organismo ha absorbido radiación ionizante y, en caso afirmativo, cuantificarla.

Una de las técnicas que ofrece la dosimetría biológica consiste en el recuento de cromosomas dicéntricos en placas metafásicas de linfocitos. Los linfocitos, o glóbulos blancos, son células de la sangre con un rol clave en la respuesta inmunitaria y son altamente sensibles a la radiación.

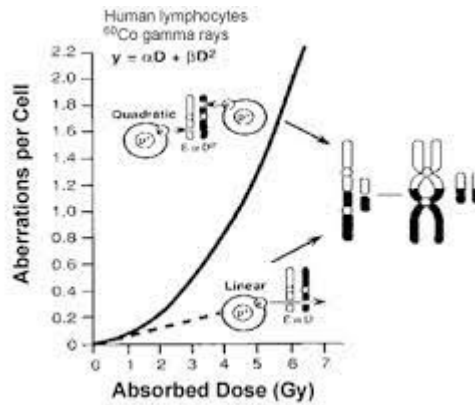


Figura 10: Frecuencia de dicéntricos en función de la dosis absorbida [17]

Se entiende por frecuencia de dicéntricos a la fracción de cromosomas dicéntricos sobre la cantidad total de cromosomas. Existen estudios que han demostrado que existe una relación entre la cantidad de radiación absorbida y la frecuencia de dicéntricos [17].

Como se puede observar en la Figura 10, la frecuencia de dicéntricos está estrechamente relacionada con la dosis absorbida:

$$Y = \alpha D + \beta D^2 + c$$

Donde Y corresponde con la frecuencia de dicéntricos, D con la dosis de radiación y c con la frecuencia de dicéntricos previa a la exposición.  $\alpha$  y  $\beta$  son coeficientes obtenidos del análisis estadístico.

Los cromosomas dicéntricos aparecen en cantidades ínfimas, prácticamente despreciables, en individuos que no han sido expuestos a radiación, por lo que el término c es despreciable. Por otra parte, el coeficiente  $\beta$  suele tender a 0 para radiaciones altamente ionizantes (objetivo de este estudio). Si se asumen estas suposiciones, la relación dosis-frecuencia de dicéntricos es lineal:

$$Y = \alpha D$$

Es por ello que, a partir de la frecuencia de dicéntricos, se puede estimar fácilmente la radiación absorbida y, por tanto, establecer el pronóstico del paciente [18]:

- Un individuo que haya recibido una dosis absorbida inferior a 2 Gy tiene altas posibilidades de sobrevivir.

- Un individuo que haya recibido una dosis absorbida comprendida entre 2 y 6 Gy tiene posibilidades de sobrevivir si recibe tratamiento adecuado.
- Un individuo que haya recibido una dosis absorbida superior a 6 Gy tiene escasas posibilidades de sobrevivir y por ende se le debe suministrar tratamiento paliativo.

# 3

## Contexto técnico

### 3.1. Estado del Arte

#### 3.1.1. Aprendizaje Profundo

En unos inicios, las redes neuronales se especializaban en una única tarea: la clasificación de imágenes. Para ello, el modelo requería de un conjunto de datos etiquetado, por lo que se trataba de un tipo de aprendizaje supervisado. El entrenamiento estaba condicionado por el volumen de datos disponible.

Más adelante, surgieron técnicas que permitieron aumentar el volumen del conjunto de datos, conocidas como *Data Augmentation*. Estas técnicas permiten conseguir nuevas imágenes a partir de la aplicación de transformaciones sobre las imágenes originales (inversiones, generación de ruido, variación de color, etc). Como resultado, se obtienen nuevas imágenes con la misma etiqueta que la imagen de la que provienen (véase la Figura 11). De esta forma, los modelos de inteligencia artificial dejarían de estar tan condicionados por el número de imágenes del conjunto inicial de datos para evolucionar a un aprendizaje auto-supervisado.

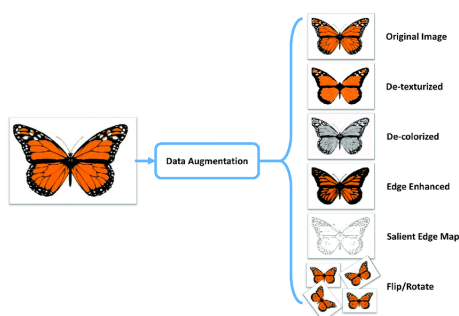


Figura 11: Data Augmentation [19]. Se observan las múltiples transformaciones que se pueden realizar sobre el conjunto de imágenes para obtener nuevas imágenes ya etiquetadas

Por otra parte, también es posible entrenar a la red neuronal sin un conjunto de datos

etiquetado (aprendizaje no supervisado). Este es el caso del aprendizaje por refuerzo (en inglés, *Reinforcement Learning*), donde el modelo se entrena en base a su propia experiencia mediante ensayo y error, fortaleciendo las conexiones que derivaron en aciertos y penalizando aquellas que derivaron en fallos.

Asimismo, también se han desarrollado redes neuronales capaces de procesar información secuencial (redes neuronales recurrentes) e incluso modelos donde una red neuronal genera contenido mientras que otra lo evalúa (redes generativas adversarias).

Todos estos avances han permitido el desarrollo de herramientas capaces de procesar el lenguaje natural humano y elaborar una respuesta (*chatbots* como ChatGPT o Gemini), reconocer el mensaje en audios de voz, e incluso generar imágenes y videos. En el ámbito de la Biomedicina, el programa AlphaFold desarrollado por DeepMind ha conseguido precedir la estructura terciaria de proteínas con alto grado de precisión [20].

### 3.1.2. Visión por Computador

Los avances en Vision por Computador han permitido la restauración de imágenes, identificación de vehículos en movimiento y avances en conducción autónoma, entre otros.

Por otra parte, la compañía Ultralytics ha desarrollado *software* que hace hoy en día posible que un modelo pueda realizar tareas diversas tanto en imágenes estáticas como en vídeo (véase la figura 12):

- **Clasificación:** Clasificación de imágenes en clases.
- **Detección:** Localización de objetos y clasificación en sus respectivas clases. Los objetos se señalan a partir del mínimo rectángulo delimitador (*bounding box*).
- **Segmentación:** Localización de objetos y clasificación en sus respectivas clases. Los objetos se señalan en todo su contorno.
- **Rastreo:** Localización de objetos en vídeos y clasificación en sus respectivas clases. Los objetos se señalan a partir del mínimo rectángulo delimitador (*bounding box*) y su movimiento se registra a partir de una línea.
- **Registro de postura:** Identificación de puntos específicos en objetos que se relacionan con partes específicas de los mismos.

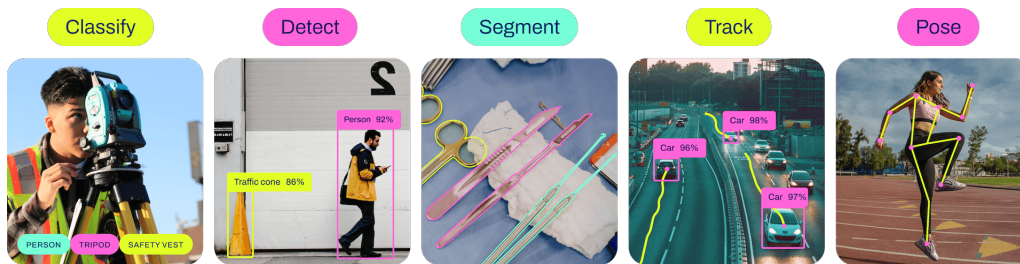


Figura 12: Múltiples tareas soportadas por Yolo v8 [21]

### 3.2. Proceso de Aprendizaje

La capacidad de aprender de un modelo de inteligencia artificial reside en la estructura computacional empleada a tal efecto. En este trabajo se han empleado redes neuronales convolucionales.

En términos generales, las redes neuronales están formadas por múltiples capas de neuronas conectadas entre sí, tal y como se ilustra en la Figura 13. Las neuronas pueden unirse a todas las neuronas de la capa posterior y previa (redes totalmente conectadas, o *fully connected*, por su nombre en inglés) o conectarse sólo con un número determinado de vecinas.

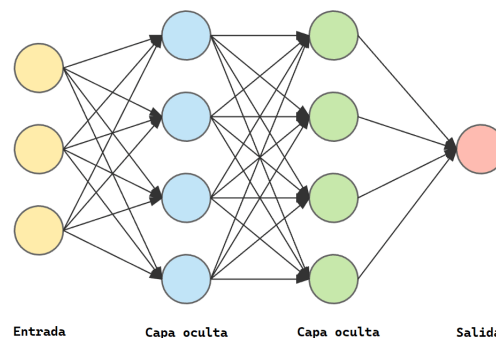


Figura 13: Arquitectura de red neuronal [22]. Nótese las distintas capas de neuronas y las conexiones entre estas.

La capa de entrada es en un vector o una matriz que representa una o varias imágenes, respectivamente. Usualmente, el conjunto de datos no se trabaja imagen a imagen, sino por lotes de imágenes (en inglés, *batches*). De ahí el uso de la matriz, que se puede comprender como un vector de varios vectores. En cualquier caso, cada uno de los elementos del vector (o de las filas de la matriz) se corresponde con cada uno de los píxeles de una imagen.

Las capas intermedias o capas ocultas son las encargadas de identificar los elementos que

permiten clasificar las imágenes. Cada capa de neuronas supone un mayor nivel de abstracción con respecto de la anterior. Así, las sucesivas capas son capaces de reconocer patrones de datos de mayor complejidad.

Al detectar el patrón que la neurona puede reconocer, la neurona se activa. En caso de no aparecer en la imagen el patrón en el que se especializa, la neurona no se activa y envía un valor nulo a las neuronas de las siguientes capas. De este modo, las sucesivas capas de neuronas pueden reconocer patrones cada vez más complejos empleando la información obtenida de capas de neuronas anteriores.

Estos patrones no son determinados por el programador, sino por la máquina. De hecho, el aprendizaje de la máquina consiste en modificar las conexiones de las neuronas con tal de encontrar los patrones adecuados para poder clasificar las imágenes correctamente.

Por otra parte, la capa de salida devuelve la predicción para la imagen. Cada una de las neuronas de esta capa representa una de las clases a predecir, siendo la predicción lanzada la correspondiente a la neurona con un valor mayor.

Para lograr predicciones correctas, el modelo necesita primero ser entrenado con un conjunto de imágenes etiquetado durante múltiples iteraciones. Durante cada iteración, conocida como época (en inglés, *epoch*) se modifican las conexiones entre neuronas –en función de la correspondencia entre la predicción y la etiqueta– mediante un proceso conocido como entrenamiento. Si el entrenamiento se realiza correctamente, el modelo es capaz de realizar predicciones correctas.

El proceso de entrenamiento consta de dos fases que se repiten por una serie de épocas. Dichas fases son la pasada hacia adelante (en inglés, *pass forward*) y la retropropagación (en inglés, *backpropagation*).

### **3.2.1. Pasada hacia adelante**

La pasada hacia adelante consiste en obtener una predicción, acertada o no, a partir de una imagen. Debido a que la imagen está previamente etiquetada, se puede conocer el error cometido por el modelo.

Se ha comentado que las neuronas se relacionan entre sí mediante conexiones neuronales. Dichas conexiones se denominan pesos, y suponen la ponderación de cada entrada que una neurona recibe. Por otra parte, cada neurona presenta un valor de sesgo, que es un valor que

aumenta o disminuye la salida de la neurona (lo que puede determinar que esta neurona se active o no).

Por tanto, la salida de cada neurona se puede hallar como sigue:

$$z = b + \sum x_i w_i$$

Siendo  $z$  la salida de una neurona cualquiera de una determinada capa,  $x_i$  el valor de cada una de las neuronas de la capa previa con las que se conecta,  $w_i$  el peso asociado a cada conexión y  $b$  el sesgo asociado a la neurona (véase la Figura 15). Cabe mencionar que la salida de las neuronas será también una de las entradas de las neuronas de la capa posterior.

Debido a que las salidas de las neuronas son una función lineal de la entrada, se ha de aplicar una función de activación  $a = f(z)$  que evite que, en todas las capas, las neuronas incrementen indiscriminadamente su valor. Por tanto, la función de activación permite establecer un umbral en la salida por debajo del cual la neurona no se activa. La función de activación  $f$  más comúnmente empleada es ReLU (*Rectified Linear Unit*).

Tal y como ilustra la Figura 14, ReLU anula la salida de la neurona si esta es negativa, y conserva su valor si esta es positiva, esto es,  $a = ReLU(z) = \max(0, z)$ .

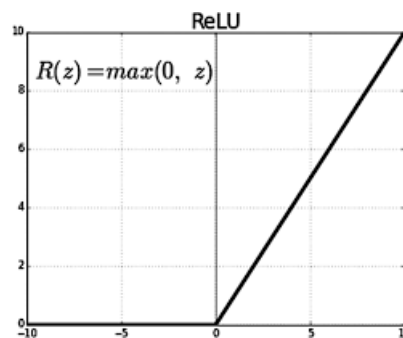


Figura 14: Función ReLU [23]. Nótese que valores negativos de la salida anulan el valor arrojado por la neurona (neurona inactiva).

A partir del valor de la salida de las neuronas de la última capa, se calcula la probabilidad de que la imagen pertenezca a la clase asociada a cada una de estas neuronas. Para ello, se emplea la función Softmax, que arrojará un valor  $\hat{y}_k$  comprendido entre 0 y 1 (el subíndice  $k$  refiere a cuál de estas neuronas se está refiriendo):

$$\hat{y}_k = \frac{e^{z_k}}{\sum e^{z_i}}$$

La clase de la neurona de salida con  $\hat{y}$  más alta (que se denotará por  $\hat{y}_n$ ) corresponderá con la predicción arrojada por el modelo.

Considerando que la imagen con la que se ha entrenado el modelo estaba previamente etiquetada, se podrá calcular el error  $\mathcal{L}_n$  cometido en la predicción:

$$\mathcal{L}_n = -\ln \hat{y}_n$$

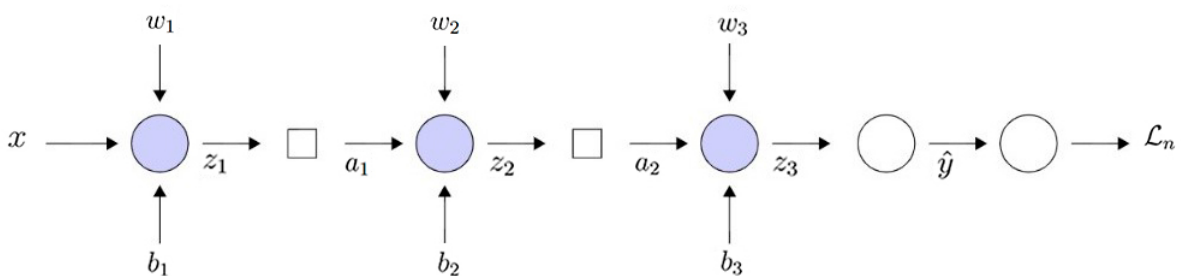


Figura 15: Pasada hacia adelante en tres neuronas cualesquiera de capas distintas y sucesivas (círculos en celeste). Los cuadrados representan funciones de activación, mientras que los dos círculos en blanco del final representan la función *Softmax* y la función de pérdida, respectivamente.

### 3.2.2. Retropropagación

La retropropagación consiste en un proceso de optimización de los pesos y sesgos de cada neurona con el fin de minimizar la función de pérdida  $\mathcal{L}$ , encontrándose el modelo en un punto  $\mathcal{L}_n$  correspondiente con el error obtenido en la pasada hacia adelante. Este proceso supone la actualización de los pesos y sesgos con valores mejores.

La función de pérdida es una función matemática multivariable, puesto que depende de cada uno de los pesos y sesgos de las neuronas del modelo. Además, la herramienta matemática que permite la optimización de funciones es la derivada.

La derivada parcial de la función de pérdida respecto de uno de sus parámetros cualesquiera en el punto correspondiente al error cometido se corresponde con el volumen tangente a la

función de pérdida en dicho punto.

Este volumen tangente corta a la función de pérdida en algún punto inferior. En dicho punto inferior el error cometido es menor, por lo que es deseable que el modelo cambie su configuración de parámetros hacia la que permite alcanzar dicho punto inferior.

Por tanto, la derivada parcial de la función de pérdida con respecto de cada parámetro permite determinar el tamaño de paso y la actualización cada parámetro.

De este modo, con cada época se podrán alcanzar progresivamente puntos de la función de pérdida con menor error. Idealmente, al final del entrenamiento se habrá alcanzado un mínimo local de la función de pérdida. Este proceso, conocido como descenso del gradiente (en inglés, *gradient descent*), se ilustra en la Figura 16.

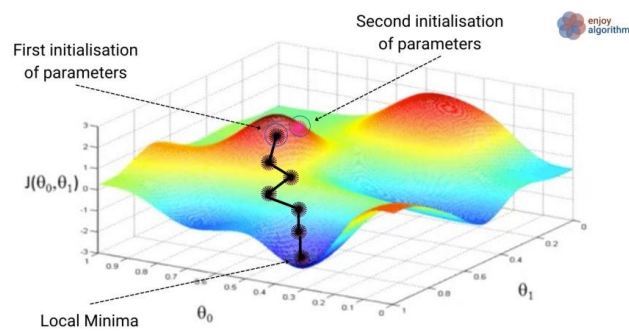


Figura 16: Descenso del gradiente [24]. Cada punto supone un menor error alcanzado en cada época hasta lograr un mínimo local. Nótese que la función de la figura es una simplificación de la función de pérdida, pues depende de tan solo dos parámetros  $\theta_0$  y  $\theta_1$ , mientras que  $\mathcal{L}$  depende de los múltiples parámetros de la red.

Se ha dicho que la función de pérdida depende de todos los pesos y sesgos de la red. No obstante, realmente la función de pérdida no depende directamente de estos parámetros, sino de las salidas de la capa neuronal previa  $\mathcal{L} = \mathcal{L}(z_1, z_2, z_3, \dots, z_n)$ .

Asimismo, cada una de estas salidas a su vez dependen de los sesgos y pesos de sus respectivas neuronas, así como de las funciones de activación de las neuronas de la capa previa  $z = z(b, a_1, w_1, a_2, w_2, \dots, a_n, w_n)$ . Por su parte, las funciones de activación dependen de la salida lineal de sus respectivas neuronas  $a = a(z)$  y, de nuevo, con estas salidas lineales ocurre que  $z = z(b, a_1, w_1, a_2, w_2, \dots, a_n, w_n)$ . Se podría continuar sucesivamente hasta llegar hasta la capa de entrada de la red, tal y como se ilustra en la Figura 17.

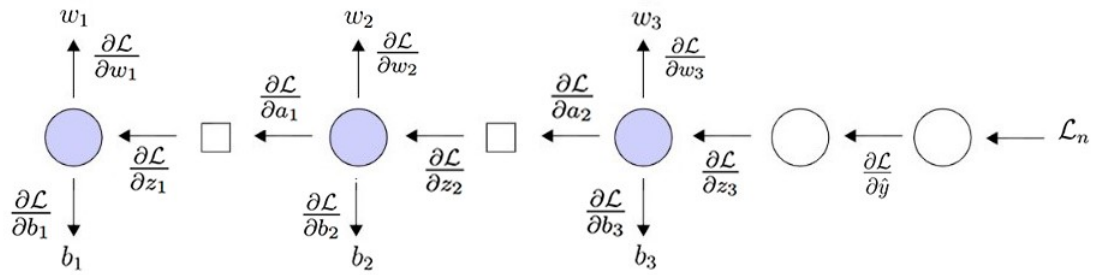


Figura 17: Proceso de retropropagación en tres neuronas cualesquiera de capas distintas y sucesivas (círculos en celeste). Los cuadrados representan funciones de activación, mientras que los dos círculos en blanco del final representan la función *Softmax* y la función de pérdida, respectivamente.

Así, los pesos y sesgos no se encuentran directamente disponibles para poder derivar la función de pérdida respecto de dichos parámetros, sino que están anidados en otras funciones. La forma de acceder a estos la proporciona la regla de derivación de la cadena, que permite propagar el error hacia todos los nodos de la red.

En primer lugar, se parte de las neuronas de la última capa, donde su salida repercute directamente en la función de error de la siguiente manera:

$$\frac{\partial \mathcal{L}}{\partial z} = \hat{y} - y$$

Una vez hallada la derivada parcial  $\frac{\partial \mathcal{L}}{\partial z}$ , se puede calcular la repercusión de la función de activación de neuronas de la capa previa:

$$\frac{\partial \mathcal{L}}{\partial a_i} = \frac{\partial \mathcal{L}}{\partial z_{i+1}} * \frac{\partial z_{i+1}}{\partial a_i}$$

En esta ocasión, se han utilizado los subíndices  $i$  para referir a neuronas de una capa y de la capa posterior. A partir de  $\frac{\partial \mathcal{L}}{\partial a_i}$  se puede hallar  $\frac{\partial \mathcal{L}}{\partial z_i}$ :

$$\frac{\partial \mathcal{L}}{\partial z_i} = \frac{\partial \mathcal{L}}{\partial a_i} * \frac{\partial a_i}{\partial z_i}$$

Por último, una vez hallado  $\frac{\partial \mathcal{L}}{\partial z_i}$ , ya sí se puede determinar cómo afecta cada parámetro a la función de pérdida  $\mathcal{L}$ :

$$\frac{\partial \mathcal{L}}{\partial w_i} = \frac{\partial \mathcal{L}}{\partial z_i} * \frac{\partial z_i}{\partial w_i}$$

$$\frac{\partial \mathcal{L}}{\partial b_i} = \frac{\partial \mathcal{L}}{\partial z_i} * \frac{\partial z_i}{\partial b_i}$$

Habiendo logrado estos valores después de propagar hacia atrás el gradiente de la función de error, se pueden actualizar los parámetros con valores más óptimos:

$$w_{t+1} = w_t - \alpha \frac{\partial \mathcal{L}}{\partial w_t}$$

$$b_{t+1} = b_t - \alpha \frac{\partial \mathcal{L}}{\partial b_t}$$

Siendo  $b_t$  y  $w_t$  los parámetros de cualquier capa empleados en la pasada hacia adelante y  $b_{t+1}$  y  $w_{t+1}$  los parámetros actualizados tras la retropropagación.  $\alpha$  es la tasa de aprendizaje (en inglés, *Learning Rate*), un hiperparámetro que multiplica al tamaño de paso y cuyo rol en el modelo se detallará en el epígrafe 'Hiperparámetros'.

### 3.2.3. Red neuronal convolucional

Una red neuronal convolucional (véase la Figura 18) es una arquitectura basada en dos redes neuronales conectadas en serie [25]: la primera se encarga de reducir las dimensiones de la imagen mediante convoluciones hasta lograr un vector que sea representativo de la imagen. La segunda toma dicho vector como entrada para realizar la predicción de la imagen, tal y como ya se ha descrito en epígrafes anteriores. Además, es importante no perder de vista que la pasada hacia adelante y la retropropagación se dan secuencialmente en ambas redes neuronales.

Este epígrafe se centra en la obtención del vector representativo de la imagen (esto es, en la primera red neuronal de las dos mencionadas). Para lograr este vector se requiere lograr un alto grado de abstracción y disminuir las dimensiones de la imagen. La convolución permite ambas.

La convolución (ilustrada en la Figura 19) es una transformación de la imagen obtenida como el producto punto de un filtro con los píxeles de la imagen con los que se superpone [25].

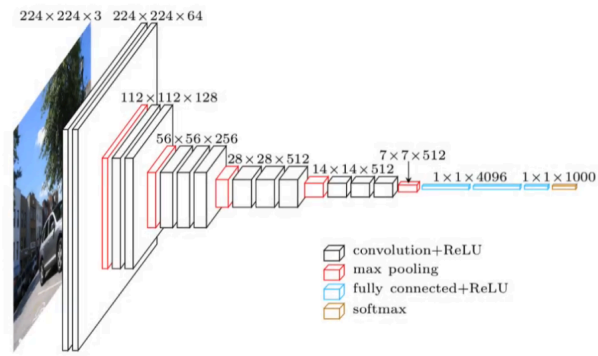


Figura 18: Estructura de red neuronal convolucional [26]. Se puede observar como la primera red disminuye las dimensiones (alto y ancho) de la imagen a la vez que aumenta el grado de abstracción (la cantidad de mapas de activación –o lo que es lo mismo, la profundidad–) a través de sucesivas capas de convolución y agrupación. Finalmente, el vector resultante entra en la red encargada de realizar la predicción (red *fully connected*).

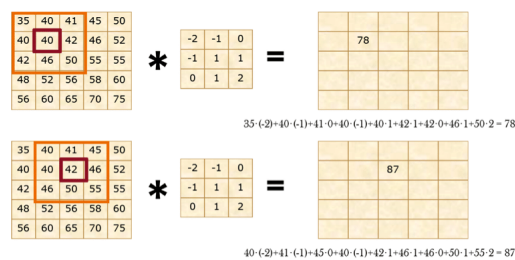


Figura 19: Proceso de convolución [27]. Se muestra el producto punto en los primeros dos píxeles de la imagen recorridos por el filtro. Al realizar dicha operación en todos los píxeles, se obtiene un mapa de activación como resultado de la convolución.

El filtro se desplaza sobre la imagen hasta haber realizado dicho producto punto sobre todos los píxeles. Como resultado, se obtiene un mapa de activación.

Los mapas de activación suponen grados de abstracción de la imagen, y actúan como neuronas que se activan o no se activan con la presencia de determinadas formas o patrones. Así, las primeras capas de neuronas son capaces de detectar determinadas combinaciones de píxeles mientras que las últimas son capaces de reconocer patrones de mayor complejidad (véase la Figura 20).

Además, una capa neuronal poseerá tantos mapas de activación como la cantidad de filtros que se hayan aplicado sobre la capa anterior.

Por otra parte, cabe mencionar que cada uno de los píxeles de los filtros son pesos que se

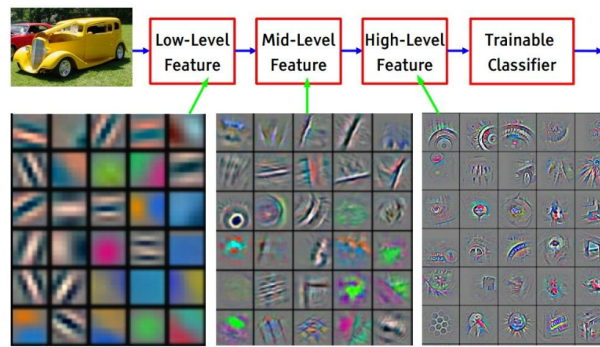


Figura 20: Mapas de activación de tres capas de una red neuronal [28]. Nótese que las capas más avanzadas reconocen patrones más complejos.

actualizan durante el proceso de entrenamiento. Según cómo se actualicen los pesos, los filtros pueden adquirir tareas como detección de bordes, afilado, enfoque, desenfoco, etc...

Esto permite al modelo encontrar cuáles son los patrones de la imagen que permiten clasificarla correctamente.

Los mapas de activación de los filtros deben reducir el tamaño (ancho y alto) de la matriz de entrada. En los casos de que la reducción de dimensionalidad no sea lo suficientemente considerable, se insertarán capas de agrupación (en inglés, *pooling*). La agrupación consiste en la reducción de un conjunto de píxeles de la imagen a un único valor, siendo el método más común la obtención del máximo valor de los píxeles sobre los que un marco se superpone (en inglés se conoce como *max pooling*). El marco se desplaza sobre la imagen hasta haber realizado la agrupación sobre todos los píxeles.

### 3.2.4. Hiperparámetros

Los hiperparámetros son configuraciones que afectan a la estructura de la red y el proceso de aprendizaje. Son previamente establecidos por el programador, por lo que una correcta elección de los mismos es decisiva para el rendimiento del algoritmo. A continuación, se enumeran los hiperparámetros más importantes.

- Tasa de Aprendizaje: Valor por el que se multiplica el tamaño de paso de cada parámetro en el descenso del gradiente. Valores acertados para dicho hiperparámetro permiten alcanzar el mínimo local más rápido, mientras que valores muy elevados pueden llegar a sobrepasar dicho punto en cada iteración, estancando al algoritmo sin llegar a reducir

el error (véase la figura 21a).

- Épocas: Número de ciclos que toma el proceso de aprendizaje. Un número reducido de épocas no será suficiente para que el modelo aprenda a reconocer los patrones en las imágenes, mientras que un número elevado de épocas puede resultar en sobreajuste, esto es, que el modelo pierda la capacidad de abstracción y solo sea capaz de clasificar correctamente las imágenes empleadas en entrenamiento, presentando un bajo rendimiento en imágenes ajenas a este subconjunto (véase la figura 21b).
- Tamaño de lote: Número de imágenes del lote con las que se entrena al modelo en cada época. Un número elevado de este hiperparámetro requiere mayor espacio en memoria, mientras que un número reducido requerirá mayor número de épocas (y mayor tiempo computacional) para poder aprovechar todas las imágenes reservadas para entrenamiento.

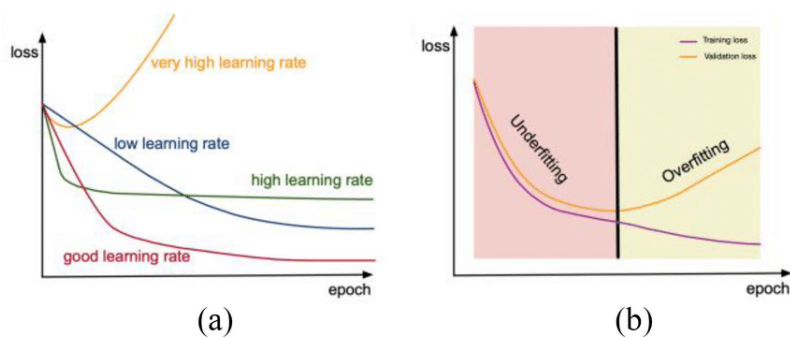


Figura 21: Las métricas obtenidas en validación permiten analizar el rendimiento del modelo con los hiperparámetros elegidos [29]. En la Figura a se observa que un acertada tasa de aprendizaje permitirá reducir el error de forma eficiente a través de las épocas. En la Figura b se observa que, si el modelo presenta sobreajuste, el error de validación se dispara mientras que el de entrenamiento se reduce.

### 3.3. Proceso de Visión

Hasta ahora se ha comentado cómo un modelo puede aprender a clasificar una imagen. No obstante, el problema que se enfrenta en este estudio requiere, no solo de un modelo que clasifique, sino que también se encargue de localizar los elementos que aparecen en la imagen.

En este contexto, localizar significa devolver el mínimo recuadro que encierra al elemento (en inglés, *bounding box*).

Debido al carácter supervisado de este tipo de aprendizaje, el *bounding box* de la predicción se puede comparar con el de la etiqueta, que se sabe que es la localización real del cromosoma (en inglés, *ground truth*). Lo deseable es que el modelo atine en las coordenadas, lo cual puede cuantificarse.

La intersección sobre la unión (IoU, del inglés, *Intersection over Union*) es un valor comprendido entre 0 y 1 que denota cuán coincidentes son el *bounding box* generado por el modelo y el *bounding box* especificado en la etiqueta de la imagen. Se obtiene como el cociente entre el número de píxeles coincidentes y el número de píxeles total que ocupan ambos recuadros, tal y como se ilustra en la Figura 22:

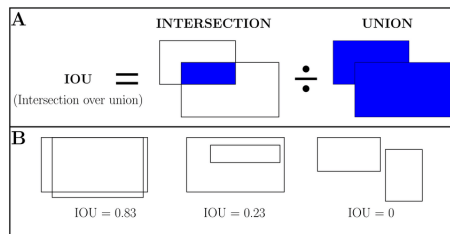


Figura 22: Intersection over Union [30]. (a) Cociente entre el número de píxeles coincidentes y el número de píxeles total que ocupan ambos recuadros. (b) Ejemplos de IoU para tres disposiciones distintas de las *bounding boxes* de la predicción y *ground truth*.



# 4

## Resultados y Discusión

### 4.1. Métricas

Para cada una de las clases que el modelo debe aprender a identificar, existen en las imágenes instancias que pertenecen a la clase (positivos) e instancias que no pertenecen a la clase (negativos). El modelo realiza predicciones para las múltiples instancias, que pueden ser correctas (verdaderas) o erróneas (falsas). Se distinguen, por tanto, cuatro tipos de predicciones desde la perspectiva de cada una de las clases:

Un verdadero positivo (TP) es una predicción para la clase que se está evaluando que se corresponde con su clase real; un falso positivo (FP) es una predicción para la clase que se está evaluando que no se corresponde con su clase real; un falso negativo (FN) es una predicción para una clase distinta a la que se está evaluando cuando esta era su clase real; un verdadero negativo (TN) se corresponde con predicción para una clase distinta a la que se está evaluando cuando esta no era su clase real.

Para poder evaluar el rendimiento del modelo se pueden emplear distintas métricas: la precisión, la exhaustividad, la sensibilidad y la especificidad. Este estudio centra su atención en las dos primeras.

Por un lado, la precisión (PR, del inglés, *precision*) es una medida que cuantifica las predicciones para una determinada clase que han sido correctas frente a las que no lo fueron. Se calcula como el cociente entre el total de verdaderos positivos y la suma de verdaderos

positivos y de falsos positivos.

$$Precision = \frac{TP}{TP + FP}$$

Por otro lado, la *exhaustividad* (RC, del inglés, *recall*) es una medida que cuantifica las instancias de una determinada clase que han sido predichas frente a las que no fueron predichas. Se calcula como el cociente entre el total de verdaderos positivos y la suma de verdaderos positivos y falsos negativos.

$$Exhaustividad = \frac{TP}{TP + FN}$$

Una medida adicional, la F-Medida, puede ser empleada para evaluar conjuntamente los valores de precisión y exhaustividad.

$$F - Medida = 2 * \frac{PR * RC}{PR + RC}$$

Además, se emplearán la exactitud y exactitud espacial. La primera cuantifica las predicciones verdaderas frente las falsas, mientras que la segunda cuantifica la cantidad de verdaderos positivos frente a falsos positivos, falsos negativos y verdaderos positivos:

$$Exactitud = \frac{TP + TN}{TP + FP + TN + FN}$$

$$Exactitud - espacial = \frac{TP}{TP + FP + FN}$$

Para cada una de las clases del modelo, se puede representar la precisión frente a la exhaustividad en un gráfico conocido como curva de precisión-exhaustividad. Al área encerrada bajo dicha función se le conoce como precisión promedio (en inglés, *average precision*) y su valor es un indicador del rendimiento del modelo, puesto que, cuanto mayor sea el área, mayor será el ratio de verdaderos positivos tanto frente a falsos positivos como a falsos negativos.

Para calcular el área bajo la curva se ha de sumar el valor de la precisión en cada uno de los puntos equidistantes en los que se divida la exhaustividad (once puntos, comúnmente) y dividir el total por el número de puntos:

$$AP = \frac{1}{11} \sum_{i=0}^{11} precision(i)$$

Cabe mencionar que los verdaderos positivos cuyo IoU sea deficiente se clasificarán, no como verdaderos positivos, sino como falsos positivos (de esta forma, se fuerza al modelo a atinar no solo en las clasificaciones, sino también en el *bounding box*). Es por ello que se debe establecer un umbral de IoU. La notación  $AP@umbral$  refiere a dicho umbral expresado en tantos por cien.

Una vez abordados estos conceptos se pueden comprender las distintas herramientas empleadas en este estudio que permiten comprobar si la actualización de los pesos y sesgos se realiza de forma adecuada:

- Función de pérdida: un gráfico muestra cómo evoluciona el error cometido por el modelo a través de las épocas. Idealmente, el error debe decrecer progresivamente hasta alcanzar un error aceptablemente pequeño. Asimismo, el decrecimiento del error de validación debe acompañar el decrecimiento del error de entrenamiento (en otro caso, si el error de validación se dispara, el modelo presenta sobreajuste).
- Matriz de confusión: un gráfico que permite visualizar el número de predicciones realizadas por el modelo frente a la clase real a la que pertenecen cada una de las predicciones realizadas. Usualmente, se suele representar en forma de gráfico donde los valores de cada celda se representan con una intensidad de color (un color más intenso representa un valor mayor). Idealmente, la matriz de confusión debe presentar valores considerablemente más altos en su diagonal principal (lo cual implica mayor número de TP frente a FP y FN).
- Media de las precisiones promedio (mAP): un gráfico muestra cómo evoluciona la media de las AP de cada clase (mAP, del inglés, *mean average precision*) a través de las épocas. Idealmente, mAP debe aumentar progresivamente hasta alcanzar un valor aceptablemente alto. Los modelos de *YOLO* empleados en este trabajo utilizan mAP@50 (umbral para IoU de 0.5).

$$mAP = \frac{1}{n} \sum_{i=1}^n AP_i$$

## 4.2. Preprocesado de Imágenes

Para la obtención de resultados, se han preprocesado las imágenes antes de entrenar al modelo con el fin de facilitar la tarea de este y, en consecuencia, obtener mejores resultados. Se han realizado múltiples experimentos en los que se han analizado los resultados de aplicar diversos filtros morfológicos de diferentes tamaños y empleando distintas técnicas. Esto, con el fin de estudiar qué combinación de los últimos permite obtener un mejor rendimiento del modelo. A continuación se procederá a comentar cada uno de dichos experimentos realizados en el preprocesado de imágenes.

Antes que nada, todas las imágenes han sido reconvertidas a escala de grises (reducción del número de canales de color de tres –RGB– a uno solo). A partir de aquí, se ha aplicado una binarización de la imagen, que consiste en tomar la intensidad de cada uno de los píxeles de la imagen (que se corresponde con un valor de gris codificado entre 0 y 256, siendo 0 un color negro y 256 un color blanco) y convertir a negro aquellos píxeles con intensidad inferior a un valor umbral y a blanco aquellos píxeles con intensidad superior. Se han implementado dos técnicas distintas para la selección de dicho valor umbral, que se detallarán en el epígrafe ‘Técnicas de umbralización’. Posteriormente, a la imagen binarizada se le han aplicado diversos filtros morfológicos y estudiado la idoneidad de cada uno, tal y como se detallará en el epígrafe ‘Filtros morfológicos’.

### 4.2.1. Técnicas de Umbralización

Para la selección del valor umbral se han empleado dos técnicas distintas, una de ideación propia y otra aplicada en un artículo también enfocado en la detección de cromosomas dicéntricos a través de una red convolucional de dos etapas que se centra en un único tipo de preprocesado de imagen [14].

La técnica de ideación propia ha consistido en obtener los histogramas de las imágenes (gráficos que manifiestan la cantidad de píxeles presenta para cada valor de intensidad) y trazar un polinomio representativo (de aquí en adelante, curva *Spline*) del conjunto discreto de puntos a través de interpolación. Sobre dicha curva, se han hallado los puntos que suponen máximos locales y se ha establecido un umbral en el valor de intensidad más próximo (hacia la ‘derecha’) en el que la cantidad de píxeles no supera la cantidad de 5 (véase la Figura 23).

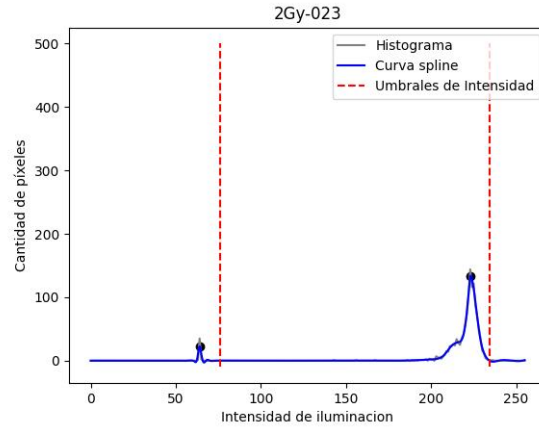


Figura 23: Histograma de la imagen 2Gy-023.jpg con curva *Spline* y umbrales obtenidos a partir de los máximos locales (pico de negros a la izquierda y pico de blancos a la derecha).

En la mayoría de casos, aparecen dos umbrales. El primero de ellos aparece a partir del pico de negros (generado por los cromosomas), mientras que el segundo aparece a partir del pico de blancos (generado por el fondo). La binarización se ha realizado utilizando el umbral que aparece a partir del pico de negros.

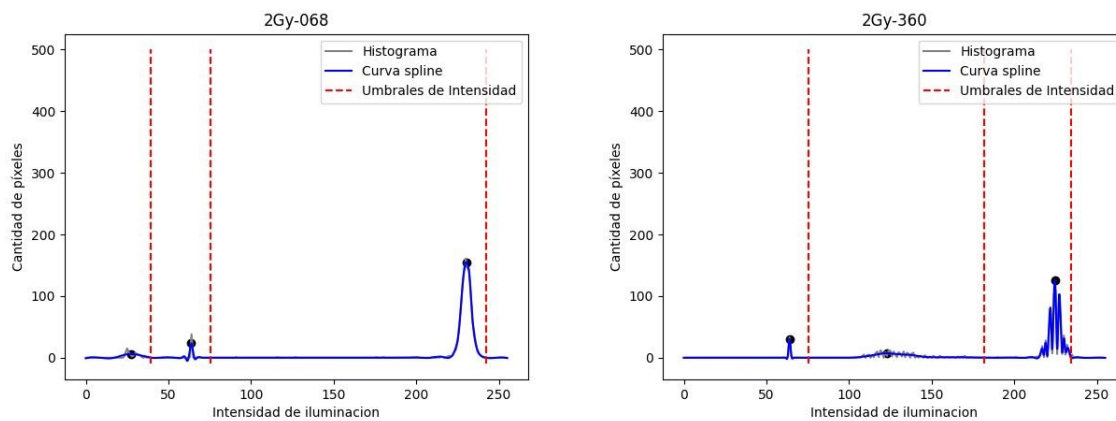


Figura 24: Histograma de las imágenes 2Gy-068.jpg y 2Gy-360.jpg con tres picos y tres umbrales.

Se han reportado dos casos (2Gy-068.jpg y 2Gy-360.jpg) en los que han aparecido tres máximos locales, y por tanto, tres umbrales (véase la Figura 24). En ambos casos, dos de esos umbrales se correspondieron con picos de negros, y además, solo la binarización que emplea el umbral generado a partir del pico de negros más grande (esto es, el pico con mayor cantidad

de píxeles) genera resultados aceptables (cromosomas reconocibles), con independencia de si el pico de negros menor se encuentra hacia la izquierda o hacia la derecha del pico de negros superior. Es por ello que con esta técnica se ha empleado, para todas las imágenes, la siguiente selección de umbral: umbral generado por el máximo de los picos, excluyendo el último pico (que siempre es superior en cantidad de píxeles ya que corresponde con el fondo).

Una vez aplicada la binarización, se ha observado que, si bien en gran parte de los cromosomas la umbralización permite reconocer los cromosomas perfectamente e incluso distinguir entre cromosomas dicéntricos y no dicéntricos, en algunas de las imágenes el umbral se encuentra muy desplazado hacia la izquierda y por tanto las imágenes no son aceptables para ser utilizadas para entrenar al modelo (a pesar de que el umbral se obtiene al alcanzar un valle del histograma).

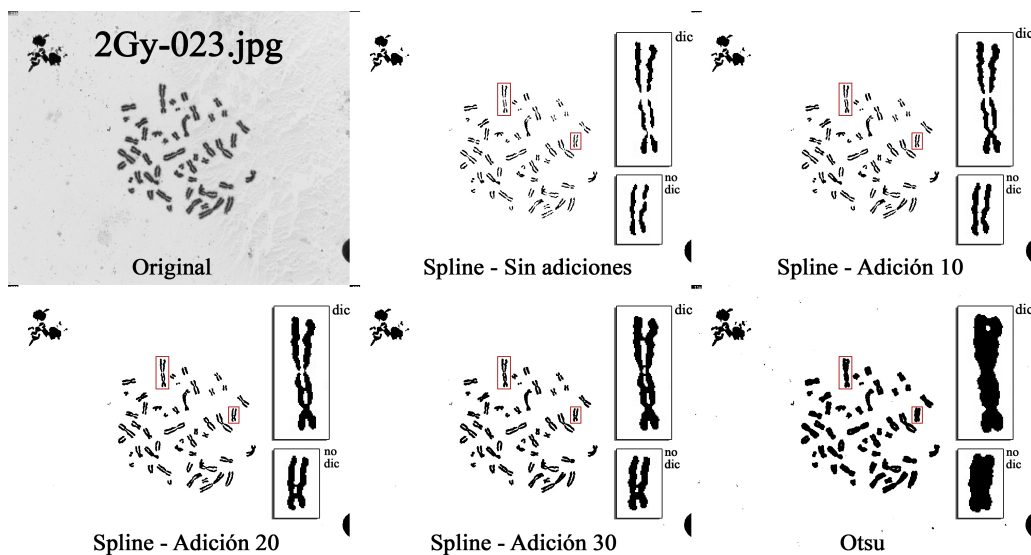


Figura 25: Comparación entre las técnicas de *Spline* con distintos valores de adición y Otsu para la imagen 2Gy-023.jpg. En este ejemplo se han mantenido constantes el filtro morfológico (apertura) y el tamaño de filtro (2x2).

Es por ello que en este método de ideación propia se ha estudiado también cómo afecta la adición de valores de 10,20 y 30 a la intensidad del umbral hallado. Si bien en algunos casos se soluciona esta situación, en otras imágenes en las que, sin adición, los cromosomas eran bien reconocibles, ahora ya no lo son. Por tanto, la idoneidad de la aplicación de adiciones reside en la particularidad de cada caso.

Por otra parte, se han estudiado los resultados obtenidos a partir de la umbralización por

el método de Otsu (empleado en el estudio anteriormente mencionado [14]). Con este método, el umbral se obtiene a partir del algoritmo desarrollado por Nobuyuki Otsu en 1979 [31].

Si bien emplear el algoritmo de Otsu para establecer el umbral soluciona el problema que aparece al emplear la técnica del Spline sin adiciones, hay que señalar que la separación entre las cromátidas hermanas es escasa, nula en muchos de los cromosomas, y que esto podría suponer una traba para el modelo a la hora de discernir entre cromosomas dicéntricos y no dicéntricos. Una comparativa entre los métodos expuestos puede apreciarse en la Figura 25.

#### 4.2.2. Filtros morfológicos

El filtro de erosión consiste en convertir en blanco todos los píxeles de la imagen que estén rodeados en cualquiera de sus ocho píxeles colindantes por un píxel con valor blanco. Por contra, el filtro de dilatación consiste en convertir en negro todos los píxeles de la imagen que estén rodeados en cualquiera de sus ocho píxeles colindantes por un píxel con valor negro.

Además, el filtro de apertura consiste en la aplicación secuencial de un filtro de erosión y un filtro de dilatación (en este orden) mientras que el filtro de cierre consiste en la aplicación secuencial de un filtro de dilatación y un filtro de erosión (en este orden). Si bien el filtro de apertura se emplea eliminar elementos salientes, el filtro de cierre se usa para unir elementos separados por un espacio estrecho.

Por otra parte, el tamaño elegido para estos filtros determinará la proporción en la que se verá afectada la imagen. En este experimento, se han aplicado filtros de apertura y cierre con tamaño de filtro de 2x2, 3x3, 4x4 y 5x5 y se han estudiado cada una de las combinaciones por separado.

Después de aplicar la binarización, se ha observado que en algunos cromosomas aparecen huecos en aquellos píxeles que en la imagen en escala de grises presentaban valores de grises más altos. Especialmente, esto ocurre en torno al centrómero y a los extremos de los cromosomas (telómeros), y podría suponer un problema ya que podría hacer desaparecer uno de los centrómeros de un cromosoma dicéntrico, pudiendo llevar al modelo a interpretar erróneamente el cromosoma como falso negativo. En estos casos, el filtro de apertura ha demostrado ser de utilidad recuperando las uniones entre las cromátidas hermanas. Sin embargo, en otros casos, el filtro de apertura provoca la aparición de puntos de cierre entre cromátidas, que no son centrómeros reales y que podría ser interpretado, por tanto, como falso positivo por el

modelo (véase la Figura 26).

Por otra parte, el filtro de cierre ha demostrado un buen desempeño en la apertura de puntos de cierre no centroméricos entre cromátidas hermanas que aparecen como producto de la binarización. No obstante, en otros casos, elimina los centrómeros del cromosoma y podría derivar en la interpretación de falsos negativos por parte del modelo (véase de nuevo la Figura 26).

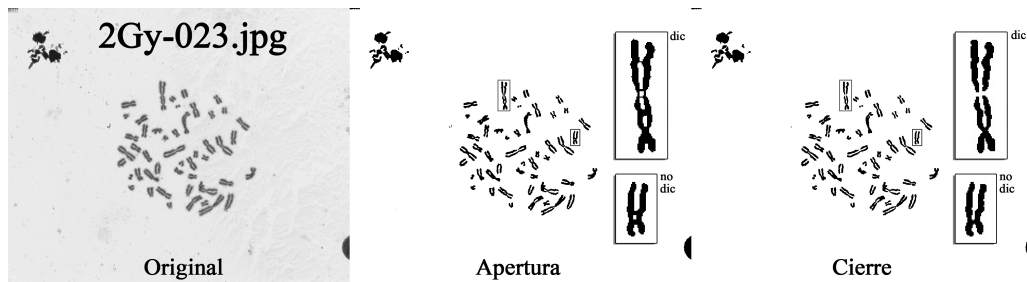


Figura 26: Comparación entre los filtros morfológicos de apertura y cierre para la imagen 2Gy-023.jpg. En este ejemplo se han mantenido constantes la técnica de umbralización (*Spline* con adición 20) y el tamaño de filtro (3x3).

Respecto al tamaño de filtro, se ha observado que los filtros de tamaño 2x2 introducen cambios muy poco notables en las imágenes, mientras que los filtros de tamaño 5x5 distorsionan enormemente los cromosomas. Los filtros de 3x3 y 4x4 introducen cambios que oscilan entre ambas situaciones extremas (véase la Figura 27).

Como se ha podido apreciar, cada una de las configuraciones introduce una serie de ventajas e inconvenientes según la imagen y los cromosomas.

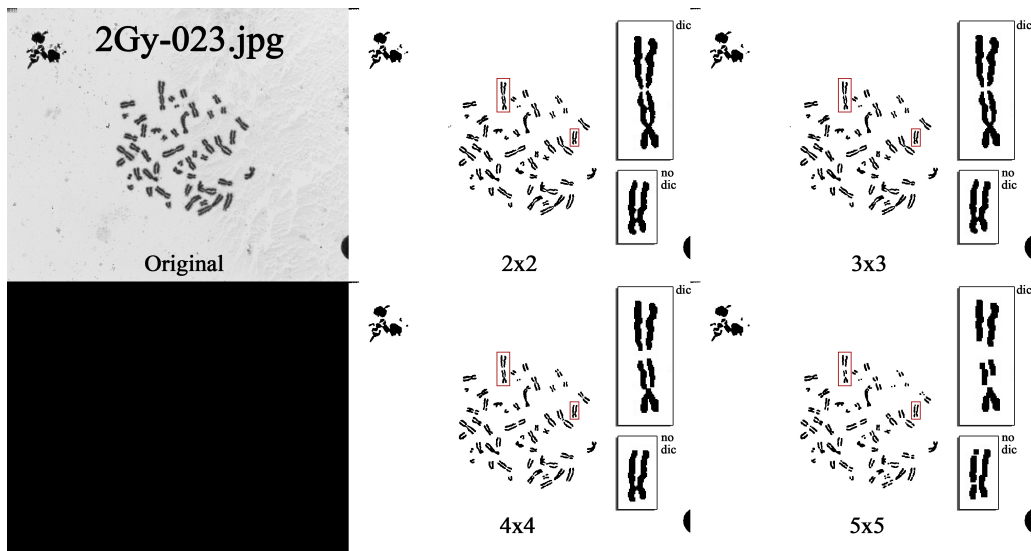


Figura 27: Comparación entre los tamaño de filtro para la imagen 2Gy-023.jpg. En este ejemplo se han mantenido constantes la técnica de umbralización (*Spline* con adición 20) y el tipo de filtro morfológico (cierre)

### 4.3. Etiquetado de imágenes

El etiquetado de dichas imágenes se ha realizado mediante la herramienta online CVAT. CVAT permite establecer el *bounding box* y descargar un archivo de texto (extensión .txt) por cada imagen, en el que cada línea representa cada uno de los cromosomas que aparecen en ella. En cada línea aparecen cinco valores separados por espacios, que se corresponden con:

1. Clase: 0 equivale a cromosoma no dicéntrico y 1 equivale a cromosoma dicéntrico.
2. Posición del centro del *bounding box* en el eje *x* (relativa respecto al ancho de la imagen).
3. Posición del centro del *bounding box* en el eje *y* (relativa respecto a la altura de la imagen).
4. Ancho del *bounding box* (relativo respecto al ancho de la imagen).
5. Altura del *bounding box* (relativa respecto a la altura de la imagen).

Cabe mencionar que las 50 imágenes han sido etiquetadas una por una de manera manual, habiéndose etiquetado un total de 2339 cromosomas, de los cuales 87 son dicéntricos y 2252

no son dicéntricos. La ratio de cromosomas dicéntricos es, por tanto, del 3.72 por ciento. En la Figura 28 se ilustra parte del proceso de etiquetado.

También es importante hacer mención de que el etiquetado ha sido posible gracias a haber disponible previamente de una copia del conjunto de imágenes donde se marcaba la localización de los cromosomas dicéntricos. Esto ha permitido que durante el proceso de etiquetado haya sido posible determinar cuáles de los cromosomas son dicéntricos y cuáles no.



Figura 28: Etiquetado de imágenes empleando la herramienta online CVAT.

#### 4.4. Validación Cruzada

La validación cruzada es una técnica de división del conjunto de datos –en subconjunto de entrenamiento y subconjunto de validación– que permite asegurar que ambas particiones son independientes entre sí.

A tal fin, para cada uno de los experimentos, el conjunto de datos (aleatoriamente reordenado) se ha partido en cinco divisiones de diez imágenes cada una y el modelo ha aprendido de forma independiente cinco veces. En cada una de estas ejecuciones, el modelo ha empleado una división distinta para validación (diez imágenes) y el resto de divisiones para entrenamiento (cuarenta imágenes).

## 4.5. Entrenamiento

Para el entrenamiento, se ha empleado la red neuronal yolov8x. Esta consta de una estructura de 640 píxeles, 68.2 millones de parámetros y presenta una velocidad en CPU (ONNX) de 479.1 ms, una velocidad en GPU A100 (TensorRT) de 3.53 ms y un rendimiento de 257.8 billones de FLOPS [32].

Debido a las limitaciones técnicas del dispositivo con el que se disponía, se han empleado los servicios en la nube de Google Colab, empleado un procesador de tipo TPU.

Asimismo, se ha entrenado y validado el modelo para cada una de las cinco combinaciones de subconjuntos de entrenamiento y validación con las imágenes generadas a partir de los siguientes experimentos: Spline apertura 2x2, Spline cierre 2x2, Spline apertura 3x3, Spline cierre 3x3, Spline apertura 4x4, Spline cierre 4x4, Spline apertura 5x5, Spline cierre 5x5, Otsu apertura 2x2, Otsu cierre 2x2, Otsu apertura 3x3, Otsu cierre 3x3, Otsu apertura 4x4, Otsu cierre 4x4, Otsu apertura 5x5, Otsu cierre 5x5. Se han seleccionado únicamente aquellas predicciones que alcanzaron al menos un valor de confianza y de intersección sobre la unión de 0.7 (con tal de evitar un elevado número de falsos positivos).

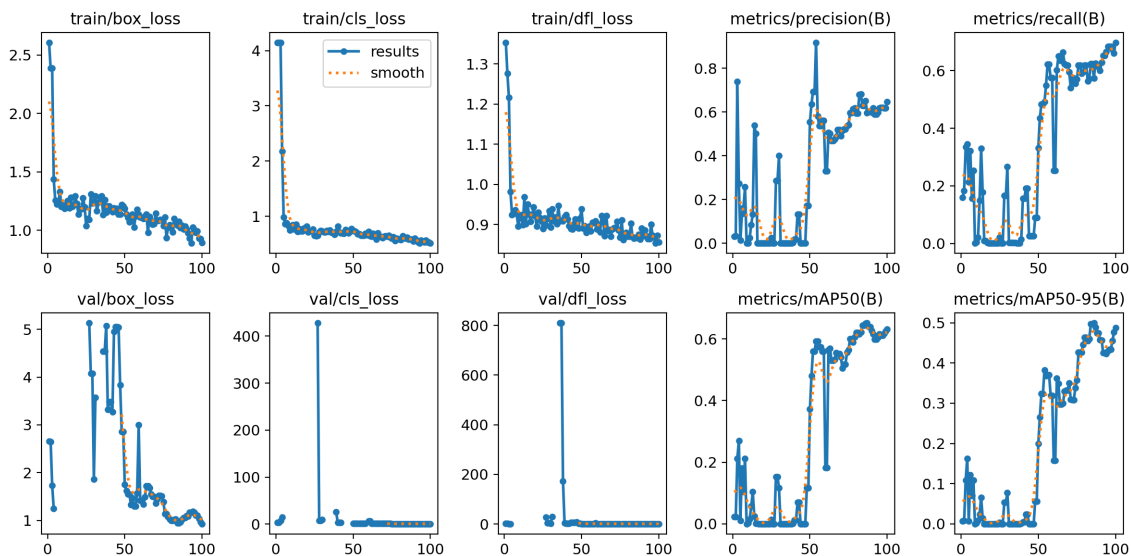


Figura 29: Rendimiento del modelo para el experimento Spline cierre 2x2 empleando la tercera subdivisión (de cinco) del conjunto de datos.

Se ha observado que, para la mayoría de subdivisiones del conjunto de datos de todos los experimentos, la función de pérdida tiende a decrecer a lo largo de las épocas (véase la

Figura 29) tanto para la detección como para la clasificación de cromosomas en dicéntricos o no dicéntricos. Este comportamiento ocurre tanto en el proceso de entrenamiento como en el de validación, por lo que el modelo no presenta características de sobreajuste. Asimismo, también cabe destacar que las funciones de precisión, exhaustividad y mAP50 tienden a crecer a lo largo de las épocas, siendo este un aspecto notable del rendimiento del modelo. No obstante, también se puede apreciar que los valores alcanzados en la última época no siempre son tan elevados como fuese deseable.

Finalmente, se han agrupado las predicciones lanzadas para los subconjuntos de validación en cada una de las ejecuciones independientes de cada experimento. De este modo, se obtuvieron las predicciones para cada una de las cincuenta imágenes en todos los experimentos. Esto permite observar el ejercicio del modelo para cada uno de los experimentos realizados.

La Figura 30, por ejemplo, se corresponde con la predicción para la imagen 2Gy-329 (Spline cierre 2x2) .

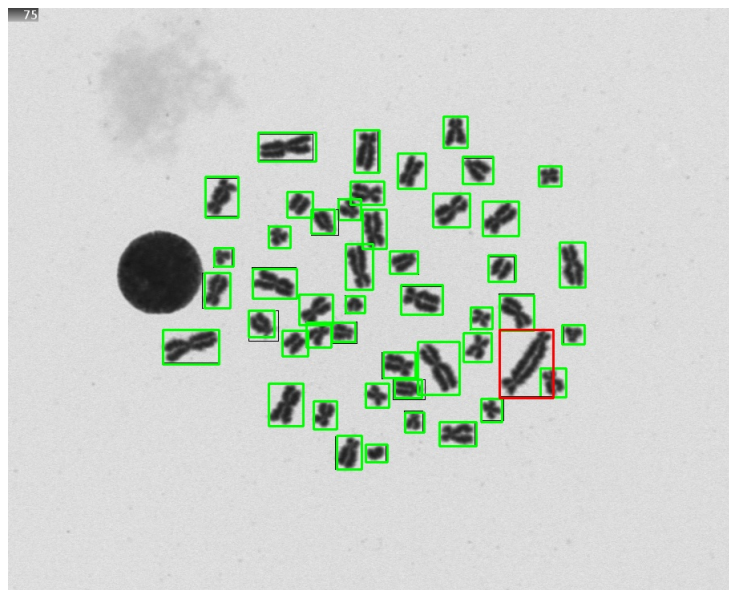


Figura 30: Predicción para los cromosomas de la imagen 2Gy-329.jpg para el experimento donde se empleó la técnica de umbralización Spline con filtro de cierre 2x2. En verde: predicciones como 'no dicéntrico'; en rojo: predicciones como 'dicéntrico'.

Desde un punto de vista cualitativo, se podría afirmar que, por norma general, los modelos no han presentado problemas en la detección de los cromosomas y que las predicciones tienen una alta correspondencia con los *bounding boxes* del *ground truth*. En lo que respecta a la

clasificación, cada experimento presenta mejores o peores resultados en función de la imagen a examinar. En las Figuras 31 y 32 se compara el desempeño de estos en el reconocimiento de un cromosoma dicéntrico (anómalo) y un cromosoma no dicéntrico (no anómalo, o normal), respectivamente.

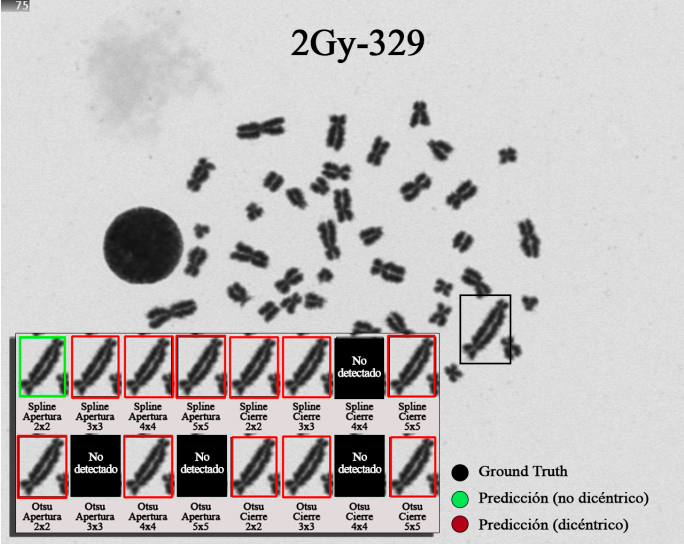


Figura 31: Composición de predicciones de la imagen 2Gy-329.jpg (para el cromosoma con índice 20) para todos los experimentos (se excluyeron adiciones). En verde: predicciones como 'no dicéntrico'; en rojo: predicciones como 'dicéntrico'.

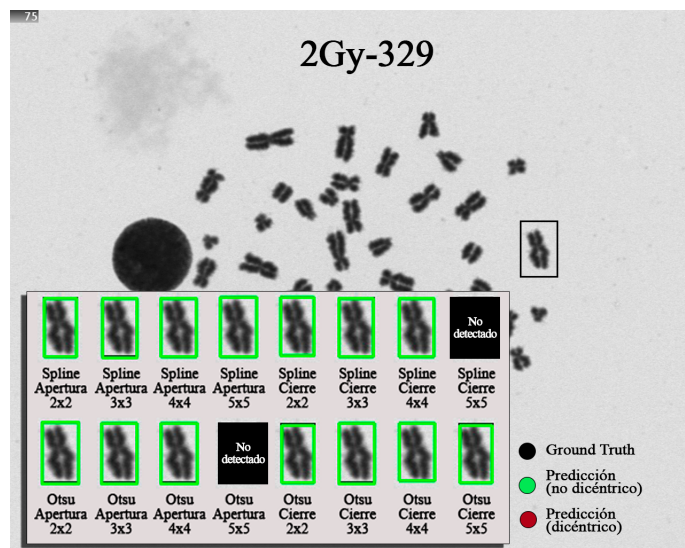


Figura 32: Composición de predicciones de la imagen 2Gy-329.jpg (para el cromosoma con índice 45) para todos los experimentos (se excluyeron adiciones). En verde: predicciones como 'no dicéntrico'; en rojo: predicciones como 'dicéntrico'.

#### 4.6. Ensamblado

Debido al gran volumen de datos generado, es importante hacer una correcta agrupación de los mismos con el fin de facilitar su análisis. A tal fin, se ha procedido a interpretar los resultados arrojados por el modelo a través de dos fases: detección y clasificación. Un esquemático de este marco de trabajo se ilustra en la Figura 33.

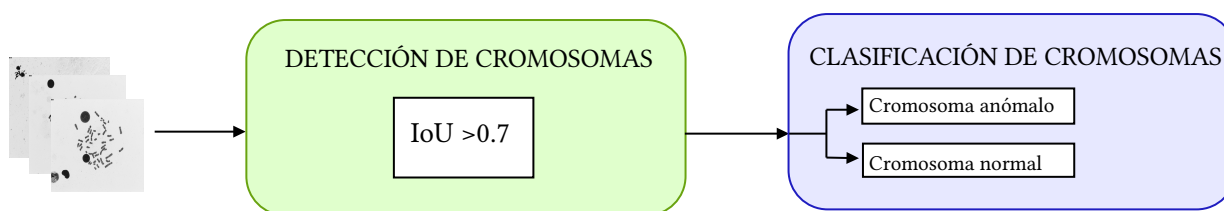


Figura 33: Fases del análisis de los datos obtenidos

En la primera de las fases, se ha estudiado si los cromosomas (*ground truth*) eran detectados por el modelo (con independencia de su clase). Para considerar que un cromosoma ha sido detectado, se ha impuesto como condición que exista una predicción que cumpla que la IoU (*Intersection over Union*) sea superior a 0.7. Posteriormente, se ha llevado un cómputo de

cromosomas detectados (verdaderos positivos), cromosomas no detectados (falsos negativos) y predicciones sin correspondencia (falsos positivos).

En la segunda fase, se ha estudiado si los cromosomas detectados eran correctamente clasificados por el modelo. Se han registrado cuatro casos: cromosomas anómalos predichos como tales (verdaderos positivos), cromosomas no anómalos predichos como tales (verdaderos negativos), cromosomas anómalos predichos como cromosomas no anómalos (falsos negativos) y cromosomas no anómalos predichos como cromosomas anómalos (falsos positivos).

Posteriormente, se ha realizado un ensamblado de datos (en inglés, *ensemble*), tal y como refleja la Figura 34, con el fin de hacer una valoración de los resultados obtenidos y determinar una función consenso que permita determinar cuál es el experimento que arroja mejores resultados. Finalmente, este preprocesado de datos puede ser seleccionado frente al resto para ser aplicado a futuras muestras que reciba el modelo para obtener una predicción.

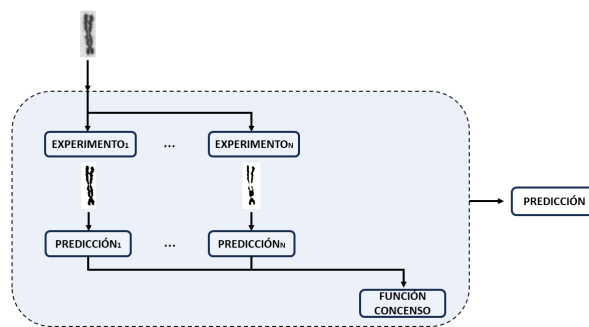


Figura 34: Ensamblado de datos.

A partir de esta información, se ha estudiado el comportamiento del modelo para cada uno de los experimentos a través de las métricas previamente expuestas en la sección 'Métricas'. Las Tablas 2 y 3 reflejan esta información.

Tabla 2: Desempeño en la detección de cromosomas dicéntricos en el total del conjunto de imágenes. La primera columna muestra cada uno de los experimentos realizados (n=Otsu, s=Spline, o=Apertura, c=Cierre, 2=2x2, 3=3x3, 4=4x4, 5=5x5). La segunda columna muestra la cantidad de predicciones obtenidas de cada tipo, mientras que la última columna supone un resumen de las métricas calculadas (en negrita, el máximo valor para cada una de ellas).

Experimento	TP	FP	FN	RC	PR	FM	S
<b>Runs nc2</b>	1909	457	430	0,816161	0,806847	0,811477	0,682761
<b>Runs nc3</b>	2075	1185	264	0,887131	0,636503	0,741204	0,58882
<b>Runs nc4</b>	1894	1909	445	0,809748	0,498028	0,616737	0,445857
<b>Runs nc5</b>	1909	457	430	0,816161	0,806847	0,811477	0,682761
<b>Runs no2</b>	2115	40	224	<b>0,904233</b>	0,981439	<b>0,941255</b>	<b>0,889029</b>
<b>Runs no3</b>	2057	28	282	0,879436	<b>0,986571</b>	0,929928	0,869033
<b>Runs no4</b>	2107	206	232	0,900812	0,910938	0,905847	0,827898
<b>Runs no5</b>	1897	3514	442	0,81103	0,350582	0,489548	0,324107
<b>Runs sc2</b>	1768	356	571	0,755879	0,832392	0,792292	0,65603
<b>Runs sc3</b>	1612	643	727	0,689183	0,714856	0,701785	0,540577
<b>Runs sc4</b>	1002	39	1337	0,428388	0,962536	0,592899	0,421362
<b>Runs sc5</b>	592	3436	1747	0,2531	0,146971	0,185959	0,102511
<b>Runs so2</b>	1704	595	635	0,728516	0,741192	0,734799	0,580777
<b>Runs so3</b>	1813	562	526	0,775118	0,763368	0,769198	0,624957
<b>Runs so4</b>	1681	780	658	0,718683	0,683056	0,700417	0,538955
<b>Runs so5</b>	1769	46	570	0,756306	0,974656	0,851709	0,741719

Tabla 3: Desempeño en la clasificación de cromosomas dicéntricos en el total del conjunto de imágenes. La primera columna muestra cada uno de los experimentos realizados (n=Otsu, s=Spline, o=Apertura, c=Cierre, 2=2x2, 3=3x3, 4=4x4, 5=5x5). La segunda columna muestra la cantidad de predicciones obtenidas de cada tipo, mientras que la última columna supone un resumen de las métricas calculadas (en negrita, el máximo valor para cada una de ellas).

Experimento	TP	TN	FP	FN	RC	PR	FM	S	ACC
<b>Runs nc2</b>	12	1847	8	42	0,222222	0,6	0,324324	0,193548	0,973808
<b>Runs nc3</b>	11	2008	9	47	0,189655	0,55	0,282051	0,164179	0,973012
<b>Runs nc4</b>	9	1792	38	55	0,140625	0,191489	0,162162	0,088235	0,950898
<b>Runs nc5</b>	12	1847	8	42	0,222222	0,6	0,324324	0,193548	0,973808
<b>Runs no2</b>	12	2043	7	53	0,184615	0,631579	0,285714	0,166667	0,971631
<b>Runs no3</b>	14	1997	2	44	0,241379	0,875	0,378378	<b>0,233333</b>	<b>0,977637</b>
<b>Runs no4</b>	12	2032	10	53	0,184615	0,545455	0,275862	0,16	0,9701
<b>Runs no5</b>	10	1834	7	46	0,178571	0,588235	0,273973	0,15873	0,972061
<b>Runs sc2</b>	6	1714	4	44	0,12	0,6	0,2	0,111111	0,972851
<b>Runs sc3</b>	5	1556	8	43	0,104167	0,384615	0,163934	0,089286	0,968362
<b>Runs sc4</b>	10	959	3	30	<b>0,25</b>	0,769231	<b>0,377358</b>	0,232558	0,967066
<b>Runs sc5</b>	5	566	4	17	0,227273	0,555556	0,322581	0,192308	0,964527
<b>Runs so2</b>	7	1629	16	52	0,118644	0,304348	0,170732	0,093333	0,960094
<b>Runs so3</b>	9	1739	16	49	0,155172	0,36	0,216867	0,121622	0,964148
<b>Runs so4</b>	8	1626	1	46	0,148148	<b>0,888889</b>	0,253968	0,145455	0,97204
<b>Runs so5</b>	11	1703	8	47	0,189655	0,578947	0,285714	0,166667	0,968909

## 4.7. Discusión

En este estudio, se han considerado la exactitud espacial (S), la exactitud (Acc) y la F-medida (Fm). Cabe recordar que estas métricas ofrecen una visión global del rendimiento del modelo y devuelven valores comprendidos entre 0 y 1 (los valores mayores son mejores).

En lo que respecta a la primera fase (detección de cromosomas), en términos generales se puede observar que el número de falsos positivos (predicciones sin correspondencia) y de falsos negativos (cromosomas sin detectar) es significativamente elevada. No obstante, existen algunos experimentos en los que la ratio de verdaderos positivos (cromosomas detectados) es lo suficientemente alta como para devolver resultados aceptables. Caben destacar los experimentos de Otsu apertura 2x2, Otsu apertura 3x3 y Otsu apertura 4x4, todos ellos con un valor de F-Medida superior al 90 por ciento. En cuanto a la exactitud espacial, solo experimentos que emplean la técnica de umbralización Otsu y filtro de apertura alcanzan un valor superior al 80 por ciento (de nuevo, Otsu apertura 2x2, Otsu apertura 3x3, Otsu apertura 4x4).

Además, resulta interesante mencionar que, en la fase de detección, el experimento Otsu apertura 2x2 es el que mejores resultados obtiene en tres métricas: exhaustividad, F-medida y exactitud espacial (0.904, 0.941 y 0.889, respectivamente).

Con respecto a la tarea de clasificación, se puede afirmar que en términos generales, los modelos arrojaron mayor número de predicciones verdaderas que falsas en los cromosomas que fueron detectados, siendo muy destacable que la mayoría de predicciones se corresponden con verdaderos negativos (no dicéntricos clasificados correctamente). No obstante, es notable que los modelos presentan dificultad caracterizando los cromosomas dicéntricos (véase la Figura 35), existiendo en todos los experimentos una significativamente mayor cantidad de falsos negativos (dicéntricos que no fueron clasificados como tales) que de verdaderos positivos (dicéntricos correctamente clasificados).

Este inconveniente queda manifiesto en las métricas de la fase de clasificación, en las que los valores para la exactitud espacial y F-medida no son aceptables para ninguno de los experimentos (los valores más altos para estas medidas no alcanzan el 40 por ciento en ningún experimento). No obstante, todos los experimentos obtuvieron una exactitud superior al 90 por ciento, siendo el experimento Otsu apertura 2x2 con el mejor valor para esta métrica (0.978).

Estos resultados pueden deberse a la reducida tasa de cromosomas dicéntricos frente a

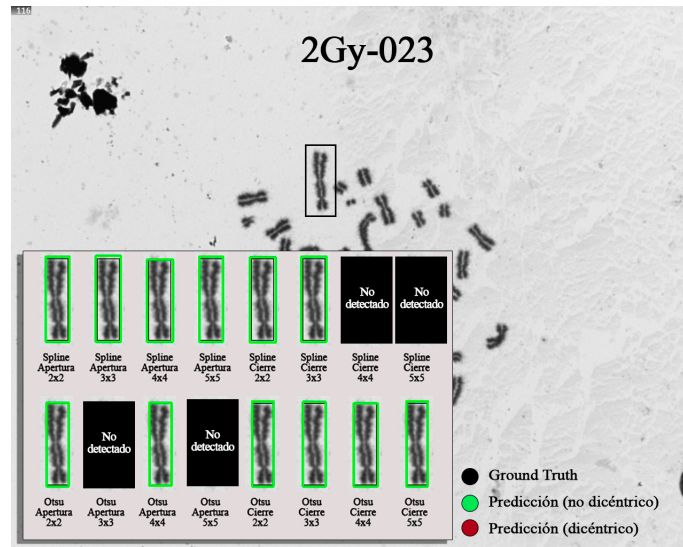


Figura 35: Composición de predicciones de la imagen 2Gy-023.jpg (para el cromosoma dicéntrico con índice 39) para todos los experimentos (se excluyeron adiciones). En verde: predicciones como 'no dicéntrico'; en rojo: predicciones como 'dicéntrico'. Nótese que este cromosoma dicéntrico pasó desapercibido en todos los experimentos

cromosomas no dicéntricos en las imágenes, siendo este un escollo para el modelo en la tarea de extraer los patrones que permiten identificar a los cromosomas dicéntricos. Asimismo, si el conjunto de datos de partida hubiese contando con un mayor volumen de imágenes o las imágenes presentaran una mayor resolución, quizá el modelo hubiese tenido mejor desempeño en el reconocimiento de los patrones que permiten clasificar a los cromosomas dicéntricos. Asimismo, cabe destacar que esta tarea ya de por sí supone todo un reto para los profesionales de la Salud, siendo mínimos los detalles que, por ejemplo, permiten distinguir un cromosoma dicéntrico de otro no dicéntrico en el que, simplemente, las cromátidas aparecen superpuestas. Este hecho también puede haber sido influyente en los resultados obtenidos.

Por otra parte, no se han observado diferencias notables en términos generales entre el rendimiento que presentan los experimentos que emplearon la técnica de umbralización Spline frente a los que utilizaron la técnica de umbralización Otsu.

Además, también se teorizó anteriormente que los filtros de apertura podrían suponer un aumento del número de falsos positivos, mientras que los filtros de cierre podría suponer un aumento de falsos negativos.

Se ha observado que, en los experimentos que aplicaron la técnica de umbralización Otsu,

el número de falsos negativos es ligeramente superior en los experimentos que aplicaron cierre frente a los que aplicaron apertura, pero no se observaron diferencias significativas en el número de falsos positivos entre filtros de apertura y cierre. Asimismo, en los experimentos que aplicaron la técnica de umbralización Spline, el número de falsos positivos es ligeramente superior en los experimentos que aplicaron apertura frente a los que aplicaron cierre, pero no se observaron diferencias significativas en el número de falsos negativos entre filtros de apertura y cierre.

# 5

## Conclusiones y Líneas Futuras

### 5.1. Conclusiones

La aplicación de técnicas de inteligencia artificial en la detección de cromosomas dicéntricos supone un enfoque prometedor para dosimetría biológica. Se trata de un campo de estudio abierto poco estudiado hasta la fecha que promete ser un gran avance en la agilización y automatización de procesos en Radiobiología y Protección Radiológica.

Este trabajo se ha centrado en el uso de redes neuronales convolucionales y visión por computador para poder detectar los cromosomas dicéntricos en micrografías de placas metafásicas de linfocitos. Para ello, se emplearon las imágenes cedidas por el centro *Radiation Emergency Assistance / Training Site (REAC/TS)* del Instituto *Oak Ridge Institute for Science and Education (ORISE)* de Tennessee, Estados Unidos.

Previamente a la fase de entrenamiento, las imágenes fueron etiquetadas manualmente empleando la herramienta online CVAT. Posteriormente, se preprocesaron las imágenes con experimentos que empleaban diferentes técnicas de umbralización, filtros morfológicos y tamaños de filtros y se entrenó el modelo con cada uno de los experimentos.

Luego, se procedió a aplicar técnicas de ensamblado de datos (nunca antes aplicadas en detección de cromosomas dicéntricos) para estudiar el rendimiento del modelo para cada experimento por separado a partir el recuento de cada tipo de predicción (TP,TN,FP,FN). A partir de estos datos, se ha podido examinar el desempeño para cada experimento tanto en el proceso de detección de cromosomas como en el de clasificación en dicéntricos o no dicéntricos.

El modelo ha seguido un adecuado proceso de aprendizaje en la mayoría de experimentos, observándose una reducción de la función de pérdida, un aumento de las métricas de exhaus-

tividad y precisión, y una ausencia de sobreajuste a lo largo de las épocas de entrenamiento.

Además, se ha observado que el modelo (con independencia del preprocesado aplicado) destaca en detectar la mayoría de cromosomas, siendo notable que los experimentos que aplican técnica de umbralización Otsu y filtros de apertura obtienen resultados significativamente mejores al resto.

En lo que respecta a la clasificación, ningún tipo de experimento destaca sobre el resto, pues en todos los casos el modelo es bueno en el reconocimiento de cromosomas no dicéntricos y malo en el reconocimiento de los dicéntricos. Una parte considerable de los cromosomas dicéntricos es inadvertida, mientras que algunos cromosomas son clasificados como dicéntricos sin serlos realmente.

Es por ello que las métricas obtenidas no fueron las deseadas, probablemente debiéndose esto a la propia naturaleza de las micrografías con una ratio dicéntricos/no dicéntricos muy baja. Cabe destacar que el reducido número de imágenes con las que se disponía, la baja resolución de las mismas y la dificultad intrínseca de la tarea puedan haber contribuido también a que estos resultados no hayan sido los esperados.

## 5.2. Líneas Futuras

A raíz de los problemas anteriormente comentados, se hace patente la necesidad de refinar el trabajo. Los siguientes aspectos podrían suponer una mejora de los resultados obtenidos:

- Aplicar técnicas de *Data Augmentation* sobre los cromosomas dicéntricos que permita aumentar la ratio dicéntricos/no dicéntricos. Asimismo, también resultaría interesante aumentar el tamaño del conjunto de datos de entrada.
- Aplicar técnicas de mejora de resolución en las imágenes.
- Asignar a cada imagen del conjunto de datos una categoría de dificultad (fácil, intermedio, difícil), analizar el comportamiento del modelo para cada una de estas y determinar qué experimentos obtienen mejor desempeño para cada categoría.
- Analizar el rendimiento de cada uno de los experimentos por separado en cada una de las imágenes y aprovechar esta característica de ensamblado para el refinamiento del modelo.

- Entrenar el modelo con los experimentos que incluyen adiciones. Cabe destacar que, si bien se preprocesó el conjunto de imágenes para cada una de las adiciones propuestas (adiciones al valor de intensidad del umbral para los experimentos que emplean la técnica de umbralización Spline), debido a las limitaciones de tiempo que implica este proyecto, se optó por no entrenar el modelo con los experimentos que aplicaban adiciones (24 experimentos).
- Entrenar el modelo con diferente combinación de hiperparámetros (para cada experimento) y estudiar cuál permite mejorar los resultados obtenidos.



# Agradecimientos

Muchas gracias a mis tutores, Miguel Ángel y Miguel José, sin quienes este trabajo no hubiera podido salir adelante. Muchas gracias de corazón por vuestro apoyo y confianza en mí.

Quisiera también agradecer enormemente al centro REAC/TS del Instituto ORISE de Tennessee, EE.UU. y a la Dra. Alegría Montoro Pastor del Hospital la Fe de Valencia por la cesión del conjunto de imágenes con el que se ha trabajado.

Especialmente, al Dr. Balajee Adayabalam, no solo por compartir las imágenes, sino también por tomarse la molestia de muy amablemente marcar manualmente los dicéntricos sobre el conjunto de datos original para que en este trabajo se pudiese realizar el etiquetado.



# Referencias

- [1] CSN, “Convocatoria subvenciones i+d 2023,” [Online] Accedido el 16-Oct-2023 en <https://www.csn.es/documents/10182/2521144/Convocatoria+Subvenciones+I%2BD+CSN+2023.pdf/523bde6d-5380-dd0f-d355-d52dc3a0806d>.
- [2] Watson, J. and Crick, F., “Molecular structure of nucleic acids: A structure for deoxyribose nucleic acid,” *Nature*, vol. 171, pp. 737–738, 1953.
- [3] LibreTexts, “10.2: Estructura y función del adn,” [Online] Accedido el 16-Oct-2023 en [https://espanol.libretexts.org/Biologia/Microbiolog%C3%ADa/Microbiolog%C3%ADa\\_\(OpenStax\)/10%3A\\_Bioqu%C3%ADmica\\_del\\_Genoma/10.02%3A\\_Estructura\\_y\\_Funci%C3%B3n\\_del\\_ADN](https://espanol.libretexts.org/Biologia/Microbiolog%C3%ADa/Microbiolog%C3%ADa_(OpenStax)/10%3A_Bioqu%C3%ADmica_del_Genoma/10.02%3A_Estructura_y_Funci%C3%B3n_del_ADN).
- [4] Quizlet, “Tema 6: Ac. nucléicos,” [Online] Accedido el 17-Oct-2023 en <https://quizlet.com/599671843/tema-6-ac-nucleicos-flash-cards/>.
- [5] FutureLearn, “Grammatical rules for dna sequence representation,” [Online] Accedido el 16-Oct-2023 en <https://www.futurelearn.com/info/courses/bacterial-genomes-bioinformatics/0/steps/47002>.
- [6] GineQualitas, “¿qué es la información genética?” [Online] Accedido el 16-Oct-2023 en <https://ginequalitas.com/informacion-al-paciente/genetica-prenatal/296-que-es-la-informacion-genetica.html>.
- [7] ResearchGate, “Esquema del ciclo celular de un organismo diploide,” [Online] Accedido el 16-Oct-2023 en [https://www.researchgate.net/figure/Figura-13-Esquema-del-ciclo-celular-de-un-organismo-diploide-se-considera-que-la-fase\\_fig8\\_319478487](https://www.researchgate.net/figure/Figura-13-Esquema-del-ciclo-celular-de-un-organismo-diploide-se-considera-que-la-fase_fig8_319478487).
- [8] Freepik, “Imágenes de mitosis,” [Online] Accedido el 16-Oct-2023 en <https://www.freepik.es/fotos-vectores-gratis/mitosis>.
- [9] Yang J. *et al.*, “Atm and atr: Sensing dna damage.” *World J Gastroenterol*, vol. 10, no. 2, pp. 155–160, 2004.

- [10] Novusmed, “¿qué es la radiación?” [Online] Accedido el 16-Oct-2023 en <https://www.novusmed.cl/acerca-de-la-radiacion/>.
- [11] A. A. F. Beatriz López-Díaz, Silvia Mercado-Sáenz and Miguel J. Ruiz-Gómez, “Genomic dna damage induced by co-exposure to dna damaging agents and pulsed magnetic field,” *International Journal of Radiation Biology*, vol. 99, no. 5, pp. 853–865, 2023.
- [12] A. F. A. Silvia Mercado-Sáenz, Beatriz López-Díaz and Miguel J. Ruiz-Gómez, “Exposure of *s. cerevisiae* to pulsed magnetic field during chronological aging could induce genomic dna damage,” *International Journal of Environmental Health Research*, vol. 32, no. 8, pp. 1756–1767, 2022.
- [13] Puerta-Ortiz, A. and Morales-Aramburo, J., “Efectos biológicos de las radiaciones ionizantes.” *Revista Colombiana de Cardiología*, vol. 27, pp. 61–71, 2020.
- [14] L. e. Shen, X., Ma, T., “High-precision automatic identification method for dicentric chromosome images using two-stage convolutional neural network,” *Nature: Sci Rep*, vol. 13, no. 2124, 2023.
- [15] T. Z. Xiang, S., Yafeng, Q., “A dicentric chromosome identification method based on clustering and watershed algorithm,” *Nature: Sci Rep*, vol. 9, no. 2285, 2019.
- [16] CSN, “Efectos biológicos de las radiaciones ionizantes,” [Online] Accedido el 17-Oct-2023 en [https://csn.ciemat.es/MDCSN/recursos/ficheros\\_md/1581136598\\_1572009112950.pdf](https://csn.ciemat.es/MDCSN/recursos/ficheros_md/1581136598_1572009112950.pdf).
- [17] I. Castro-Volio, “Indicadores citogenéticos para la identificación de exposición a radiación ionizante en humanos.” *Acta Médica Costarricense*, vol. 55, no. 3, pp. 110–117, 2013.
- [18] M. MSD, “Lesiones causadas por la radiación,” [Online] Accedido el 16-Oct-2023 en <https://www.msmanuals.com/es-es/hogar/traumatismos-y-envenenamientos/lesiones-causadas-por-la-radiaci%C3%B3n/lesiones-causadas-por-la-radiaci%C3%B3n>.
- [19] Engati, “Data augmentation,” [Online] Accedido el 25-Oct-2023 en <https://www.engati.com/glossary/data-augmentation>.
- [20] John Jumper *et al.*, “Highly accurate protein structure prediction with alphafold.” *Nature*, vol. 596, pp. 583–589, 2021.

- [21] Ultralytics, “Ultralytics yolov8 tasks,” [Online] Accedido el 22-Oct-2023 en <https://docs.ultralytics.com/tasks/>.
- [22] P. Huet, “Qué son las redes neuronales y sus aplicaciones,” [Online] Accedido el 22-Oct-2023 en <https://openwebinars.net/blog/que-son-las-redes-neuronales-y-sus-aplicaciones/>.
- [23] I. Alvarado, “Redes neuronales,” [Online] Accedido el 26-Oct-2023 en [https://ml4a.github.io/ml4a/es/neural\\_networks/](https://ml4a.github.io/ml4a/es/neural_networks/).
- [24] R. Raj, “Gradient descent algorithm in machine learning,” [Online] Accedido el 31-Oct-2023 en <https://www.enjoyalgorithms.com/blog/parameter-learning-and-gradient-descent-in-ml>.
- [25] E. L. de Falken, “Visión artificial con redes convolucionales (cnn),” [Online] Accedido el 23-Oct-2023 en <https://www.ellaberintodefalken.com/2019/10/vision-artificial-redes-convolucionales-CNN.html>.
- [26] M. Sotaquirá, “¿qué son las redes convolucionales?” [Online] Accedido el 23-Oct-2023 en <https://www.codificandobits.com/blog/redes-convolucionales-introduccion/>.
- [27] Giménez-Palomares F. *et al.*, “Aplicación de la convolución de matrices al filtrado de imágenes.” *Modelling in Science Education and Learning*, vol. 9, p. 97, 2016.
- [28] R. Gandhi, “Build your own convolution neural network in 5 mins,” [Online] Accedido el 25-Oct-2023 en <https://towardsdatascience.com/build-your-own-convolution-neural-network-in-5-mins-4217c2cf964f>.
- [29] Maribel Torres-Velázquez *et al.*, “Application and construction of deep learning networks in medical imaging,” *IEEE Transactions on Radiation and Plasma Medical Sciences*, pp. 1–1, 2020.
- [30] Salau, J. and Krieter, J., “Instance segmentation with mask r-cnn applied to loose-housed dairy cows in a multi-camera setting,” *Animals*, vol. 10, p. 2402, 2020.
- [31] N. Otsu, “A threshold selection method from gray-level histograms.” *IEEE*, vol. 9, no. 1, pp. 62–66, 1979.

[32] Ultralytics, “Ultralytics yolov8,” [Online] Accedido el 30-Ene-2024 en <https://github.com/ultralytics/ultralytics>.



UNIVERSIDAD  
DE MÁLAGA

| [uma.es](http://uma.es)

E.T.S de Ingeniería Informática  
Bulevar Louis Pasteur, 35  
Campus de Teatinos  
29071 Málaga

E.T.S. DE INGENIERÍA INFORMÁTICA