

Máster Executive en Data Science
para Profesionales
MBIT School

DEEP LEARNING APLICADO A IMÁGENES MÉDICAS

María Teresa Tardío
Francisca Fuentes
Tommaso Giuseppini



Mayo 2020

Presentación

María Teresa Tardío



Graduada en Química por la universidad de Alcalá de Henares (UAH) con Máster en Industria Farmacéutica y Parafarmacéutica. María Teresa ha trabajado durante los últimos años en el sector farmacéutico desempeñando funciones de técnico de laboratorio/validaciones. Durante el último año ha cursado el Máster de Data Science para profesionales en Mbit School compaginandolo con su trabajo actual (junior data science, sector consultoría). [maite.tardio@gmail.com]

Francisca Fuentes



Egresada de Analista Programador Computacional por DuocUC, sede Padre Alonso Ovalle (2014 - Santiago de Chile). Francisca se ha dedicado cuatro años al sector IT como Desarrolladora Back-End y Analista de Bases de Datos. Cuenta con un bootcamp de UX/UI Design por Ironhack (2017 - Madrid). Durante los casi tres últimos años ha trabajado en el sector de Tecnología e Innovación como Diseñadora de Experiencia de Usuario. Desde hace un año cursa el Máster Executive en Data Science en MBIT School. [francisca.fuentes2811@gmail.com]

Tommaso Giuseppini



Diploma di Liceo Artístico Alessandro Caravillani y dos años de curso in Rilevamento dei beni Architettonici en la Università La Sapienza di Roma. Tommaso Giuseppini ha trabajado durante los últimos quince años, como ingeniero de soporte UNIX y soluciones high-end HW en compañía puntera como ORACLE, desarrollando elevadas competencias en el ámbito de las infraestructuras de sistemas. En los últimos dos años amplía sus conocimientos en lenguajes de programación y cursa el Máster de Data Science para profesionales en la escuela madrileña Mbit School. [tgiuseppini@gmail.com]

Agradecimientos

El presente proyecto no podría haberse realizado sin los conocimientos adquiridos durante el Máster de Data Science para profesionales de Mbit School.

La realización de este trabajo es fruto de las orientaciones, sugerencias y estímulos de los profesores Manuel Montañez y Emilio Soria, quienes nos han conducido estos meses y mostrado en cada momento una inmejorable disposición ante las dudas que nos surgieron durante la realización del mismo, aportando valiosas observaciones que, en todo momento, guiaron esta investigación.

Deseamos agradecer también la colaboración que hemos tenido con la empresa valenciana QUIBIM S.L., en especial a Rafael López, quien nos brindó la oportunidad de colaborar con ellos, así como resolvernos dudas y proporcionarnos el material necesario (dataset y documentación).

Por último y no menos importante, queremos agradecer a nuestros familiares y amigos el apoyo que nos han dado durante todo este tiempo, y a la tremenda familia que hemos ganado con nuestros compañeros de MBIT y los lazos que hemos generado este último año.

Abstract

In this project we have set ourselves the objective of facing the challenge of medicine “Image Denoising”. Diagnostic imaging is a fundamental field in both the prevention and detection of diseases, and for this reason, treatment prior to analysis by a professional becomes a phase of great importance. The purpose of this research is focused on the study of the different types of noise distributions that are most frequently observed and generated in magnetic resonance imaging and their elimination through the use of a specific model of Convolutional Neural Network (AUTOENCODER). The types of noise studied are Riciano, Salt and Pepper and Gaussian. The image dataset consisting of RMI of prostates has been provided by QUIBIM S.L.

Resumen

En este proyecto nos hemos puesto como objetivo enfrentarnos al reto del “Denoising” de imágenes en la medicina. El diagnóstico por imagen es un campo fundamental tanto en el ámbito de la prevención como en la detección de enfermedades y por esta razón el tratamiento previo a la análisis por parte de un profesional se convierte en una fase de gran importancia. El fin de esta investigación se centra en el estudio de los diferentes tipos de distribuciones de ruidos que más frecuentemente se observan y se generan en las resonancias magnéticas y su eliminación a través del uso de un modelo específico de red convolucional profunda (AUTOENCODER) . Los tipos de ruidos estudiados son el Riciano, Salt and Pepper y el Gaussiano. El dataset de imágenes que consta de MRI de próstatas ha sido proporcionado por QUIBIM S.L.

Índice general

1. Introducción	1
2. Estado del Arte	5
3. Teoría y Fundamentos en Imagen Médica	11
3.1. Modalidades de Adquisición Ionizantes:	11
3.1.1. Rayos X (RX)	11
3.1.2. Tomografía Computarizada (TC)	15
3.1.3. Medicina Nuclear (MN)	16
3.1.3.1. Tomografía por Emisión de Positrones (PET)	17
3.1.3.2. Imagen Híbrida y Nuevas Formas de Imagen	22
3.2. Modalidades de Adquisición No Ionizantes:	33
3.2.1. Resonancia Magnética (RMN)	33
3.2.2. Ultrasonido (US)	52
3.3. Estándares de Imagen Médica:	64
3.3.1. DICOM	64
3.3.2. NIFTI	75
3.3.3. HL7	80
4. Machine Learning y Deep Learning en Medicina	85
4.1. Machine Learning	85
4.2. Deep Learning	89
4.3. Redes Neuronales Profundas Convolucionales CNNs	93
4.4. Red Convolucional U-Net	105
4.5. Autoencoders	109
4.6. Tipos de Ruido	117
4.6.1. Salt and Pepper:	118
4.6.2. Gaussiano:	119
4.6.3. Riciano:	120
5. Metodología	123
5.1. Imágenes de Origen	123
5.2. Preprocesado de las Imágenes	124

5.2.1.	Imágenes Originales y Normalización de Píxeles	125
5.2.2.	Añadir Ruido y Normalización de los Datos	127
5.2.2.1.	Ruido Gaussiano	127
5.2.2.2.	Ruido Salt and Pepper	130
5.2.2.3.	Ruido Riciano	133
5.3.	Creación del Conjunto de Datos para el Entrenamiento de los Modelos y su Validación	135
5.4.	Construcción de los Modelos	136
5.5.	Métricas de Validación	139
5.5.1.	Coeficiente de Determinación (R^2)	140
5.5.2.	Coeficiente de Determinación Ajustado	140
5.6.	Herramientas y Tecnología	141
6.	Resultados	142
6.1.	Ruido Salt and Pepper	142
6.1.1.	Primer Ruido Salt and Pepper	142
6.1.1.1.	Entrenamiento de los Modelos	143
6.1.1.2.	Validación de los Modelos	147
6.1.2.	Segundo Ruido Salt and Pepper	149
6.1.2.1.	Entrenamiento de los Modelos	149
6.1.2.2.	Validación de los Modelos	153
6.2.	Ruido Gaussiano	155
6.2.1.	Primer Ruido Gaussiano	155
6.2.1.1.	Entrenamiento de los Modelos	155
6.2.1.2.	Validación de los Modelos	159
6.2.2.	Segundo Ruido Gaussiano	161
6.2.2.1.	Entrenamiento de los Modelos	161
6.2.2.2.	Validación de los Modelos	165
6.3.	Ruido Riciano	167
6.3.1.	Primer Ruido Riciano	167
6.3.1.1.	Entrenamiento de los Modelos	167
6.3.1.2.	Validación de los Modelos	171
6.3.2.	Segundo Ruido Riciano	173
6.3.2.1.	Entrenamiento de los Modelos	173
6.3.2.2.	Validación de los Modelos	177
7.	Conclusiones	180
8.	Trabajo Futuro	182
9.	Bibliografía	183

1. Introducción

Al día de hoy, la inteligencia artificial está jugando un papel crucial en la detección de enfermedades en el campo de la medicina. Se puede destacar la localización automática de órganos y su segmentación así como clasificación de imágenes.

Para un resultado óptimo de éstas aplicaciones es de vital importancia el tratamiento previo de los datos, como la reducción de las perturbaciones (*ruido*) que se pueden producir a la hora de generar las distintas imágenes. Este ruido puede influir en los resultados obtenidos por el modelo, pudiendo obtener falsos negativos, diagnóstico que en el campo de la medicina no se puede permitir.

El objetivo de este proyecto es desarrollar un modelo que sea capaz de eliminar ruido en imágenes (RMI) obteniendo como resultados una reconstrucción lo más ajustada posible a una imagen de entrada sin perturbación.

Los beneficios que pueden aportar modelos de reducción de ruido son:

- No sería necesaria la repetición de la prueba si la imagen obtenida presenta ruido. Con ello se podría evitar someter al paciente a pruebas que pueden ser perjudiciales para su salud como Rayos X, o pruebas que a pesar de no ser perjudiciales son incómodas para el paciente como puede ser la Resonancia magnética.
- Reducción de costes: las técnicas que se emplean para obtener las imágenes presentan costes elevados, poder evitar la repetición y el uso de maquinaria específica ayudaría a reducir esos costes.
- Mejora en la gestión de recursos: el uso de los equipos específicos y personal cualificado para realizar las distintas pruebas es limitado, produciéndose largas listas de espera de pacientes.

A continuación se muestra una gráfica donde se puede observar el número de exámenes por resonancia magnética (RMI) en diferentes países en el año 2017 (por cada 1.000 habitantes)[1]:

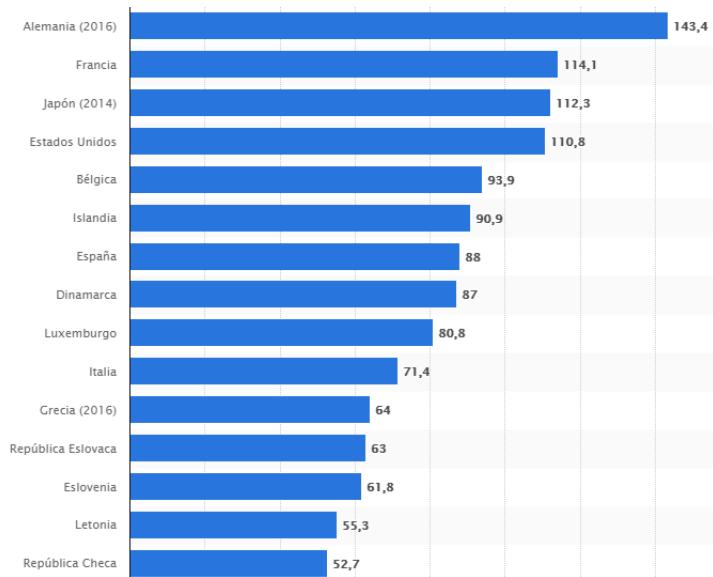


Figura 1.1: Pie de foto 1

En los siguientes gráficos se puede apreciar el tiempo de espera de los pacientes para realizar una determinada prueba médica. En primer lugar se presenta un gráfico de los distintos períodos de espera en Italia (2019) [2].



Figura 1.2: Pie de foto 2

A continuación se pueden observar listas de espera en España (2019) para la realización de resonancia magnética [3]:

LISTAS DE ESPERA EN LOS DEPARTAMENTOS DE SALUD PARA RESONANCIA MAGNÉTICA

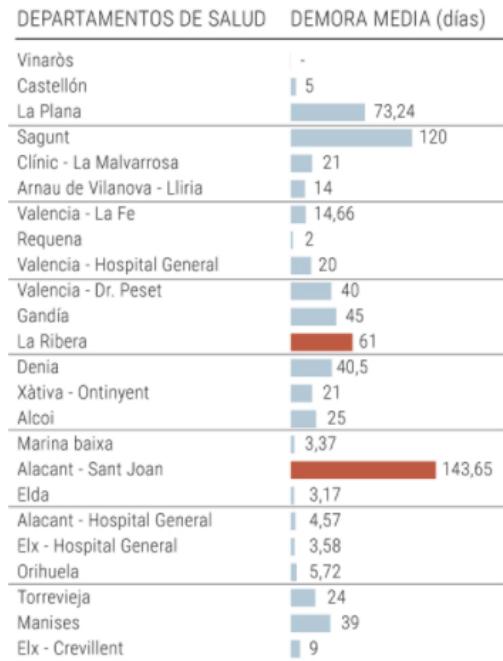


Figura 1.3: Pie de foto 3a

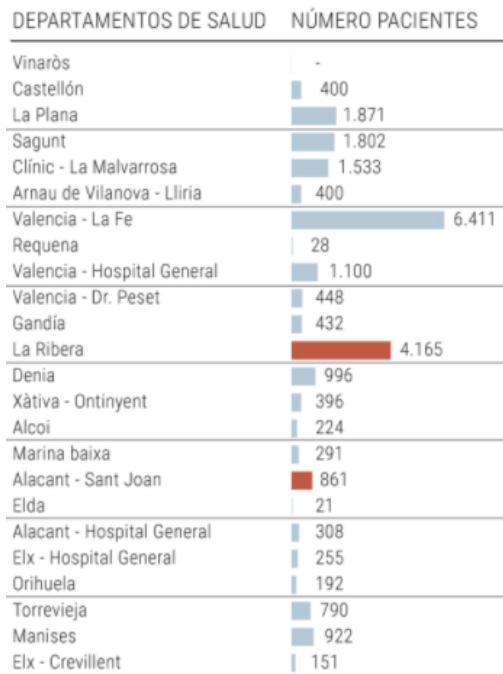


Figura 1.4: Pie de foto 3b

Como podemos observar es necesario la aplicación de técnicas que permitan la disminución de las listas de espera tan elevadas.

La utilización de modelos específicos para realizar operaciones de denoising en imágenes médicas evitaría los problemas mencionados anteriormente, ya que nos proporcionan resultados óptimos para ser utilizados por especialistas u otros modelos matemáticos.

A continuación se muestra un ejemplo de uso de uno de los modelos desarrollados durante este proyecto:

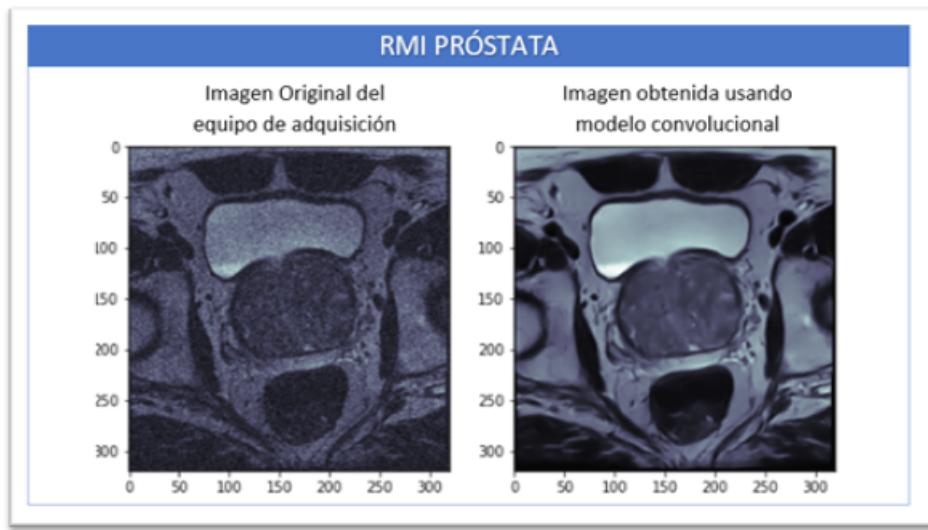


Figura 1.5: Pie de foto 4

Este trabajo de Fin de Máster entra dentro del marco de colaboración de la escuela Mbit con la empresa valenciana QUIBIM S.L. desde donde nos propusieron la realización del proyecto cuyo objetivo se enfocase a la disminución de ruido en imágenes DICOM de resonancia magnética (RM).

QUIBIM S.L. fue creada a fines de 2012 como una iniciativa de radiólogos e ingenieros biomédicos como un spin-off del Instituto de Investigación de Salud de La Fe (Valencia). QUIBIM aplica inteligencia artificial y modelos computacionales avanzados a las imágenes radiológicas para medir de manera objetiva los cambios producidos por una lesión o por un tratamiento farmacológico, ofreciendo información cuantitativa adicional al enfoque cualitativo de la radiología.

2. Estado del Arte

A Novel Convolutional Autoencoders Based Multilevel Denoising Framework for Medical Image Data

Abstract:

La eliminación de ruido de imagen basada en autoencoders convolucionales ha surgido como un enfoque prometedor en aplicaciones de imágenes médicas. La mayoría de los métodos de eliminación de ruido tradicionales, como el filtro bayesiano, el filtro de mediana no lineal, la transformación de shearlet basada en wavelets, etc., son difíciles de procesar imágenes médicas debido a la existencia de ruido aditivo, multiplicativo y gaussiano. Además, estos modelos no pueden resolver el problema de la escasez en las imágenes médicas comprimidas. Para superar estos problemas, se propone una novedosa técnica de eliminación de ruido multinivel basada en sensores de compresión para minimizar la tasa de ruido de las imágenes médicas en el marco de la red neuronal de convolución (CNN). Esta técnica se implementa en el marco de autoencoders convolucionales para reducir el ruido en cada capa de red profunda y mejorar la calidad de las imágenes. Esta técnica elimina eficazmente las imágenes médicas utilizando la varianza inter, intra en el marco de CNN. Los resultados experimentales demostraron que el marco propuesto tiene mejores medidas de PSNR y SSIM en comparación con los enfoques tradicionales de vanguardia.

[4] Kumar, K. K. (2020). A NOVEL CONVOLUTIONAL AUTOENCODERS BASED MULTILEVEL DENOISING FRAMEWORK FOR MEDICAL IMAGE DATA. *Journal of Critical Reviews*, 7(2), 30-37.

The Impact of Neural Network Techniques in the Optimization of the Image Processing

Abstract:

Hoy en día, el procesamiento de imágenes se utiliza ampliamente en muchas aplicaciones y para diversos fines. Los académicos propusieron y sugirieron varias técnicas de procesamiento de imágenes. La red neuronal es una de las principales técnicas de procesamiento, que es un método de vanguardia. Este artículo tiene como objetivo investigar técnicas de redes neuronales en el campo del procesamiento de imágenes. Además, las imágenes médicas, así como las tendencias crecientes de la utilización de imágenes médicas digitales, han ganado una gran atención en los sectores de la salud. En este sentido, este artículo se centra en el efecto de las redes neuronales en la optimización del procesamiento de imágenes médicas. En este contexto, el diagnóstico precoz y la detección del ojo tienen un papel importante en la prevención de la discapacidad visual, debido al hecho de que alrededor de 45 millones de personas tienen discapacidad visual en todo el mundo, según la Organización Mundial de la Salud. Por esta razón, el artículo actual presenta un nuevo método basado en el procesamiento de imágenes para la segmentación vascular basada en un contorno morfológico activo. Luego, la segmentación se realizó en base a operaciones morfológicas, c-medias difusas y transformación de cuenca. El resultado de tales métodos de segmentación se dio a la red neuronal convencional. Los valores de las características optimizadas se extraen. El valor umbral se establece para comparar estos valores de características optimizados. A partir de esto, se obtendrán los mejores métodos de segmentación.

[5] Anwer, D. A. (2020). The Impact of Neural Network Techniques in the Optimization of the Image Processing. *Journal of Southwest Jiaotong University*, 55(2).

Medical Image Denoising Using Convolutional Denoising Autoencoders

Abstract:

La eliminación de ruido de imágenes es un paso importante de preprocessamiento en el análisis de imágenes médicas. Se han propuesto diferentes algoritmos en las últimas tres décadas con diferentes rendimientos de eliminación de ruido. Más recientemente, después de haber superado todos los métodos convencionales, los modelos basados en el aprendizaje profundo han mostrado una gran promesa. Sin embargo, estos métodos están limitados para el requisito de un gran tamaño de muestra de capacitación y altos costos computacionales. En este artículo mostramos que con un tamaño de muestra pequeño, los autoencoders de eliminación de ruido construidos usando capas convolucionales pueden usarse para la eliminación de ruido eficiente de imágenes médicas. Las imágenes heterogéneas se pueden combinar para aumentar el tamaño de la muestra y aumentar el rendimiento de eliminación de ruido. Las redes más simples pueden reconstruir imágenes con niveles de corrupción tan altos que el ruido y la señal no son diferenciables para el ojo humano.

[6] Gondara, L. (2016, December). Medical image denoising using convolutional denoising auto-encoders. In 2016 IEEE 16th International Conference on Data Mining Workshops (ICDMW) (pp. 241-246). IEEE.

Wavelet-Based Rician Noise Removal for Magnetic Resonance Imaging

Abstract:

Es bien sabido que los datos de imagen de magnitud de resonancia magnética obedecen a una distribución de Rician. A diferencia del ruido aditivo gaussiano, el ruido "de Rician depende de la señal, y separar la señal del ruido es una tarea difícil. El ruido de Rician es especialmente problemático en los regímenes de baja relación señal / ruido (SNR) donde no solo causa fluctuaciones aleatorias, sino que también introduce un sesgo dependiente de la señal en los datos que reduce el contraste de la imagen. Este artículo estudia los métodos de filtrado del dominio wavelet para la eliminación del ruido Rician. Presentamos un nuevo filtro de dominio wavelet que se adapta a las variaciones tanto en la señal como en el ruido.

[7] Nowak, R. D. (1999). Wavelet-based Rician noise removal for magnetic resonance imaging. *IEEE Transactions on Image Processing*, 8(10), 1408-1419.

Noise in Magnitude Magnetic Resonance Images

Abstract:

Los objetivos de este artículo son revisar las propiedades del ruido en las imágenes MR de magnitud para aclarar la terminología utilizada al referirse al ruido y desalentar el uso de los términos ruido de Rician y sesgo de ruido de Rician. Se sabe que la distribución de las intensidades de píxeles MR medidas en presencia de ruido es Rician, y el ancho de este La distribución está directamente relacionada con el ruido gaussiano en las señales reales e imaginarias medidas. Son los valores de magnitud de píxel los que siguen la distribución de Rician, no el ruido. El término ruido de Rician debe usarse con precaución o, mejor aún, evitarse por completo, ya que es inherente a esta terminología el comportamiento que normalmente no está asociado con el ruido, como la dependencia de la intensidad de la señal. Esta terminología es engañosa y puede conducir a malentendidos conceptuales y prácticos. Es mejor relacionar el ruido de la imagen con el ruido gaussiano en las señales reales e imaginarias.. © 2008 Wiley Periodicals, Inc. Concepts Magn Reson Part A 32A:409–416, 2008.

[8] Cárdenas-Blanco, A., Tejos, C., Irarrazaval, P., & Cameron, I. (2008). Noise in magnitude magnetic resonance images. *Concepts in Magnetic Resonance Part A: An Educational Journal*, 32(6), 409-416.

Evaluating Denoising Performances of Fundamental Filters for T2-Weighted MRI Images

Abstract:

En las imágenes de resonancia magnética (MR), el ruido es un problema común que limita la precisión de la imagen de cualquier medición cuantitativa. La eliminación de ruido en el preprocesamiento de imágenes de MRI es un paso importante para eliminar el ruido y hacer que la imagen se ajuste a otros pasos involucrados en el proceso de análisis. Sin embargo, los diferentes tipos de ruidos producen rangos de impacto significativo en la calidad de la imagen y, por lo tanto, tienden a afectar la interpretación humana y el rendimiento de los sistemas de diagnóstico asistidos por computadora. Otro tema es acerca de las estrategias de filtrado para eliminar el ruido y preservar la imagen de alta calidad según la capacidad de reconstrucción del filtro y el modelo de ruido. En este trabajo, se utilizan tres algoritmos de filtrado diferentes, como filtro medio (MF), filtro adaptativo (ADF) y filtro medio (AVF) para eliminar los ruidos aditivos presentes en las imágenes de resonancia magnética, es decir, gaussiano, sal y ruido de pimienta y moteado. La densidad de ruido se agregó gradualmente a la imagen de IRM hasta en un 90 % para comparar el rendimiento de los filtros mediante evaluación cualitativa y cuantitativa. El rendimiento de estos filtros se compara utilizando los parámetros estadísticos como el error cuadrático medio (MSE) y la relación señal / ruido máxima (PSNR). El estudio muestra que el filtro Median reconstruye una imagen de alta calidad que otros filtros en Gaussian y en la eliminación de sal y pimienta con PSNR de 38.3 dB con una variación de ruido del 10 %. Mientras que para la eliminación del ruido moteado, el filtro promedio funciona mejor que otros, lo que resulta en 56.2 dB PSNR con una variación de ruido del 10 %. Una comparación con otros métodos bien establecidos, este estudio muestra que el filtro Mediana y Media produce mejores resultados de eliminación de ruido, preservando las estructuras y detalles principales.

[9] Isa, I. S., Sulaiman, S. N., Mustapha, M., & Darus, S. (2015). Evaluating denoising performances of fundamental filters for t2-weighted MRI images. Procedia Computer Science, 60, 760-768.

3. Teoría y Fundamentos en Imagen Médica

3.1. Modalidades de Adquisición Ionizantes:

▪ Técnicas Basadas en Radiaciones Ionizantes:

- Rayos X (RX)
- Tomografía Computarizada (TC)
- Medicina Nuclear (MN)
 - Tomografía por Emisión de Positrones (PET)
 - Imagen Híbrida y Nuevas Formas de Imagen

3.1.1. Rayos X (RX)

En el año 1895, Wilhelm Conrad Röntgen (1845-1923) científico alemán de la Universidad de Würzburg, descubrió una radiación (de origen desconocido en aquel momento, y de ahí su nombre de rayos X) que tenía la propiedad de penetrar los cuerpos opacos [10].



Figura 3.1: Descubrimiento de Rayos X

Los rayos X son invisibles a nuestros ojos, pero producen imágenes visibles cuando usamos placas fotográficas o detectores especiales para ello (Figura x, radiografía de un perro).

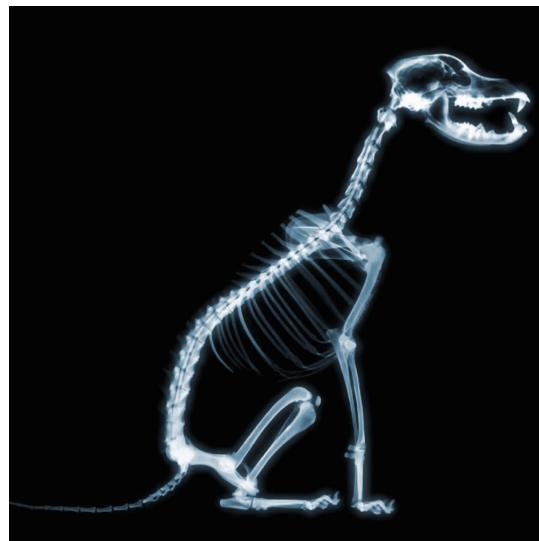


Figura 3.2: Radiografía de un perro

Actualmente existen múltiples aplicaciones de los Rayos X en el campo de la Medicina tales como: radiografías, estudio de los vasos sanguíneos (angiografía), etc.

Desde el descubrimiento de los Rayos X en 1895 hasta que la aplicación de este tipo de radiación revolucionó los campos de la Física, la Química y la Biología pasaron varios años.

Max Von Laue (1879-1960, premio Nobel de Física en 1914) era profesor en las Universidades Munich, Zurich, Frankfurt, Würzburg y Berlín en 1912 demostró la naturaleza ondulatoria de la nueva radiación mediante el análisis de la difracción de los mismos en las moléculas de los cristales cristalinos.

Para ello, colocó cristales de sulfatos de cobre y de blenda frente a los rayos X, obteniendo la confirmación de su hipótesis y demostrando al mismo tiempo la naturaleza periódica de los cristales (ver Figura 3.2).

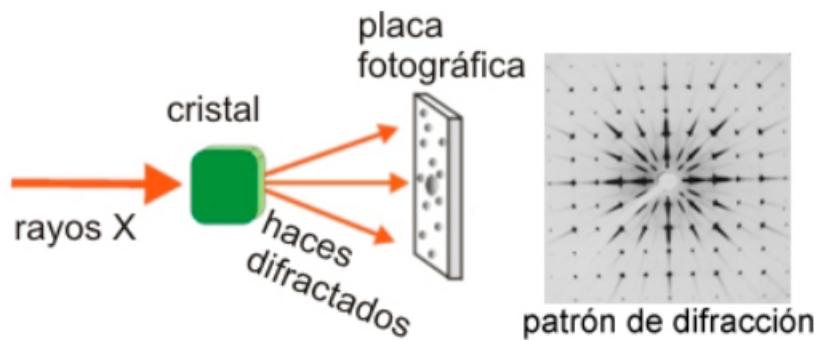


Figura 3.3: Difracción de Rayos X

Los Rayos X son una forma de radiación electromagnética al igual que la luz visible con una longitud de onda del orden de $10^{-10}m$ (1 Angstrom).

Generación de los Rayos X

Los rayos X [11] se producen en el interior de un tubo de vidrio (Figura X. Tubo de Rayos X), en el que se ha hecho un alto vacío, y donde se aplica una diferencia de potencial (d.d.p.) de aproximadamente 50 a 150 KV entre sus polos positivo (ánodo) y negativo (cátodo). El cátodo se calienta cuando hacemos pasar una corriente entre sus extremos, produciéndose una emisión de electrones por efecto termoiónico.

Los electrones emitidos por el cátodo encuentran un campo eléctrico que los atrae hacia el ánodo (polo positivo), debido a la elevada diferencia de potencial.

Los electrones se dirigen hacia el ánodo sin chocar en su trayectoria con partículas de polvo o aire, ya que el tubo se encuentra con un alto vacío, cediendo en él la energía que transportan básicamente de dos formas, por interacciones con los electrones corticales y por interacciones con los núcleos.

Las interacciones con los electrones corticales, dan lugar a choques elásticos (no hay emisión de radiación electromagnética, pero si se produce calor provocando el calentamiento del ánodo) e inelásticos (provocan la promoción de electrones a capas superiores, que al caer a sus órbitas iniciales emiten energía).

Los átomos excitados tienden a relajarse en un intervalo de tiempo muy corto. Así el electrón excitado pasa desde el nivel en el que se encuentra al fundamental, directamente o a través de niveles intermedios emitiendo en cada uno de estos saltos un fotón de energía igual a la diferencia de energía entre los niveles:

$$hv = E_1 - E_2$$

donde E_1 y E_2 representan respectivamente las energías de las órbitas entre las que se produce el salto.

Cuando el electrón excitado o emitido desde el átomo (ionización) pertenece a una capa muy profunda (K o L normalmente) la posterior vuelta al estado fundamental se realiza mediante la emisión de fotones muy energéticos denominados rayos X.

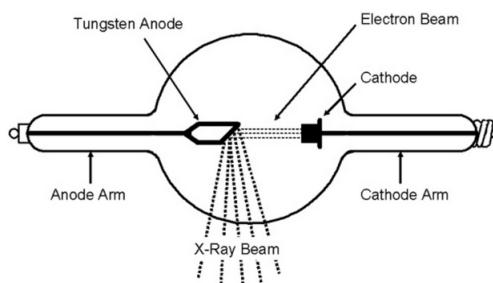


Figura 3.4: Pie de foto 8

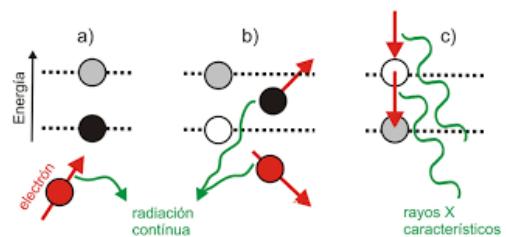


Figura 3.5: Pie de foto 9

Propiedades de los Rayos X

Este tipo de radiación tiene la capacidad de atravesar la materia Produciendo los siguientes efectos:

- **Efecto biológico:** La radiación se atenúa al atravesar la materia, lo que significa que parte de ella es absorbida, produciendo lesiones en los organismos vivos. :
- **Efecto luminiscente:** Producen fluorescencia en ciertas sustancias llamadas fósforos.
- **Efecto fotográfico:** Impresionan y producen imágenes sobre las películas fotográficas.
- **Efecto ionizante:** Pueden ionizar los gases.

3.1.2. Tomografía Computarizada (TC)

Tomografía Computarizada (TC) [12] es un procedimiento en el cual un paciente es sometido a un haz de rayos X el cual se gira rápidamente alrededor del cuerpo, produciendo señales que son procesadas para generar imágenes transversales (cortes) del cuerpo llamados imágenes tomográficas.

Una vez recogidos varios cortes sucesivos, éstos, se pueden apilar digitalmente para formar una imagen tridimensional del paciente, permitiendo la identificación y ubicación de las estructuras básicas, así como de posibles tumores o anomalías.

Funcionamiento de TC

Un escáner de TC utiliza una fuente motorizada de rayos X que gira alrededor de una abertura circular de una estructura en forma de dona (donut) llamada Gantry.

Durante el escaneo por TC, el paciente permanece tumbado en una cama que se mueve lentamente a través del Gantry, mientras que el tubo de rayos X gira alrededor del paciente, disparando haces angostos de rayos X a través del cuerpo.

Los fotones que consiguen atravesar el cuerpo son medidos por detectores situados en oposición a la fuente de rayos X y transmitidos a una computadora.

Inicialmente los equipos de TC contaban con un solo detector. En la actualidad la mayoría de equipos cuentan entre 8 y 64 filas de detectores aproximadamente 1000-2000 detectores por fila.

Una vez finalizado el proceso de adquisición, la computadora de TC utiliza técnicas matemáticas sofisticadas para reconstruir un corte de imagen 2D del paciente.

Este proceso se lleva a cabo mediante los siguientes tipos de algoritmos matemáticos:

- **Analíticos:** Uno de los más utilizados es el “Filtered Back Projection” (FBP).
- **Iterativos:** Son más complejos. Originalmente difíciles de implementar debido a que la capacidad computacional no era la adecuada.

Calidad de imágenes obtenidas en TC

En la TC existe un compromiso entre la calidad de la imagen y la dosis de radiación ionizante [9]. El problema radica en que al disminuir las dosis de radiación en TC, se aumenta la cantidad de ruido en las imágenes. Esto se debe a que los detectores del tomógrafo reciben una cantidad menor de fotones, lo que disminuye la relación señal a ruido.

De este modo, para conservar un balance adecuado, el principio por el cual se rige la TC es el de usar las menores dosis de radiación que conservan razonablemente la calidad en la imagen. En inglés, este término se conoce como ALARA (*As Low As Reasonably Achievable*).

Para disminuir el ruido en las imágenes se han propuesto diferentes algoritmos y métodos matemáticos. Entre estos destacan algoritmos que filtran directamente las proyecciones de rayos X, o las imágenes reconstruidas.

En 1995, Tomasi y Manduchi propusieron un método llamado el filtro bilateral, capaz de preservar los bordes de la imagen y disminuir el ruido en regiones uniformes.

El filtro bilateral (FB) es un filtro adaptativo no lineal.

La idea principal del FB es que cada píxel (o voxel) se procesa teniendo en cuenta un criterio de distancia espacial así como una medida de la diferencia en intensidades de gris.

Por ejemplo, dos píxeles contiguos tenderán a promediarse si se utiliza sólo un criterio de distancia espacial, similar a como lo hace un filtro Gaussiano; sin embargo, cuando se usa el FB también se tiene en cuenta la diferencia en intensidades de gris. Si la intensidad es parecida, de acuerdo con un parámetro predefinido, los píxeles tratarán de igualar sus intensidades.

Si las intensidades son muy distintas (por ejemplo en los bordes de una imagen), dicha diferencia se aplica en el cálculo y ambos píxeles preservarían sus intensidades originales.

3.1.3. Medicina Nuclear (MN)

¿Qué Estudia la Medicina Nuclear?

La medicina nuclear es una rama de la medicina que se encarga de utilizar radiotrazadores o fármacos con isótopos radiactivos para poder obtener imágenes funcionales (son capaces de reflejar el funcionamiento actual de los órganos o tejidos a estudiar) y moleculares (pueden mostrar alteraciones a un nivel molecular).

El proceso está conformado por la aplicación del fármaco, generalmente por vía intravenosa, el cual se distribuirá por todo el cuerpo o en un órgano en concreto según las necesidades de cada paciente, este paso puede llevar desde horas hasta un par de días en completarse.

Una vez en el cuerpo se podrá detectar el isótopo por medio de un detector de radiación para posteriormente procesar la información obteniendo las imágenes detalladas del sistema a tratar.

Este procedimiento se utiliza mayormente para diagnosticar aquellas enfermedades que pueden pasar desapercibidas ante otros chequeos, entre ellas se pueden encontrar; diversos tipos de cáncer, trastornos neurológicos, problemas gastrointestinales, cardíacos, musculares y muchos otras más.

3.1.3.1. Tomografía por Emisión de Positrones (PET)

La Tomografía por Emisión de Positrones (PET por su sigla en inglés) es una técnica de medicina nuclear que produce imágenes de distribución de diferentes moléculas, por lo que se considera una técnica de «imagen molecular». Proporciona información de función, que se complementa con la información de forma suministrada por técnicas como la Resonancia Magnética Nuclear (RM) o la Tomografía Axial Computarizada (TAC), para producir imágenes de gran utilidad, principalmente en oncología y, en menor medida, en cardiología, neurología y otras especialidades.

La Tomografía por Emisión de Positrones (PET por su sigla en inglés) es una técnica de medicina nuclear que produce imágenes de distribución de diferentes moléculas, por lo que se considera una técnica de «imagen molecular». Proporciona información de función, que se complementa con la información de forma suministrada por técnicas como la Resonancia Magnética Nuclear (RM) o la Tomografía Axial Computarizada (TAC), para producir imágenes de gran utilidad, principalmente en oncología y, en menor medida, en cardiología, neurología y otras especialidades.

La PET se basa en la administración endovenosa de trazadores análogos de la glucosa ligados a marcadores radioactivos emisores de positrones, que son captados por las células del organismo. Al no seguir el ciclo bioquímico de la glucosa y abstenerse de pasar al ciclo de Krebs, los trazadores quedan atrapados en la célula. La retención del trazador está influenciada directamente por la tasa metabólica celular, por lo que en condiciones normales, órganos como el cerebro, el hígado y el miocardio exhiben una retención muy pronunciada. Esto quiere decir que, el trazador se acumula en tejidos específicos del cuerpo, o mejor explicado aún, cuanto más activo sea el tejido (por ejemplo, cuanto mayor sea la cantidad de glucosa o de oxígeno que utiliza), mayor será la cantidad de trazador acumulada y mayor será la radiación que emite.

Por su mayor metabolismo, las células malignas tienen una alta afinidad por la glucosa y sus análogos, cuyo paso al interior de la célula está también potenciado. El mismo principio de acción de la PET determina sus limitantes. Los tumores de bajo grado de malignidad pueden no ser detectados dada su menor tasa metabólica y, por consiguiente, menor captación del trazador. Asimismo, no resulta sencillo distinguir entre la concentración del trazador en tejidos comprometidos por células malignas y la que se observa cuando se está ante un proceso inflamatorio o infeccioso activo. Ciertas condiciones fisiológicas, patológicas y farmacológicas afectan la captación de glucosa por los distintos órganos, lo que hace necesario conocerlas de antemano para una adecuada interpretación del estudio. A nivel encefálico, y dada la peculiaridad del cerebro, cuyo metabolismo depende casi que exclusivamente del consumo de oxígeno y glucosa, cualquier circunstancia que aumente el pool sanguíneo de glucosa (consumo de esteroides, diabetes, inobservancia de ayunos, etc.) provocará un fenómeno de competencia con el trazador y, en consecuencia, un incremento en la incidencia de falsos negativos[13].

Lo importante de esta técnica es que ofrece una información funcional (bioquímico-metabólica) de alta sensibilidad, lo que la diferencia de otras exploraciones de diagnóstico por la imagen que hacen hincapié preferentemente en los aspectos morfológicos[14].

Características

- Las modalidades de imagen de medicina nuclear difieren en gran medida de las modalidades de imagen de “transmisión”, como RM o TC.
- La fuente en el caso del PET es un radioisótopo inyectado por vía intravenosa previamente a la adquisición.
- La dosis de radiación es relativamente baja debido a la corta vida de los radioisótopos.
- Aunque su resolución espacial es baja, el PET ofrece información a nivel molecular que no es posible obtener con RM o TC.
- Algunos de los radioisótopos más comúnmente utilizados en PET son ^{18}F , ^{11}C o ^{13}N .
- La unión del radioisótopo con determinadas moléculas permite obtener radiofármacos con distintas aplicaciones.
- Por ejemplo, el uso de ^{18}F junto con glucosa genera fluorodesoxiglucosa (FDG). Aquellos tejidos ávidos de glucosa, como pueden ser aquellos afectados por procesos tumorales, captarán el radiofármaco en un grado muy superior al tejido sano.

Bases Físicas:

1. **Proceso de Aniquilación** Las principales ventajas del PET surgen de las propiedades físicas de la emisión de positrones. Cuando un positrón, electrón cargado positivamente, es emitido desde el núcleo, viaja una corta distancia perdiendo energía hasta que interactúa con un electrón del medio, de modo que ambos se aniquilan (desaparecen). La masa del electrón y del positrón se convierte en energía bajo forma de dos rayos gamma (cada uno de 511 keV), que viajan en direcciones opuestas (π 180°). La energía del positrón determina la distancia que recorre antes de la aniquilación, pero siempre el resultado de ésta es la producción de dos fotones de 511 keV. Por tanto, al contrario del SPECT donde normalmente es emitido un fotón único en cada desintegración, en PET es emitido simultáneamente un par de fotones y en consecuencia su detección involucra un par de detectores en situación opuesta que debe registrar eventos en un mismo instante de tiempo (o sea, en coincidencia)[15].
2. **Punto de Aniquilación** Debido a que dos fotones viajan en direcciones opuestas, el punto de aniquilación estará ubicado en una línea recta que une ambos puntos de detección. Esto significa que la información direccional se puede determinar “electrónicamente” sin la necesidad de una colimación convencional. Al contrario de las cámaras gamma, la detección no se limita a aquellos fotones que viajan en ángulos rectos respecto al detector y en consecuencia la sensibilidad es varias veces mayor en PET que para SPECT. La colimación se mantiene normalmente para separar datos de diferentes planos, sin embargo en cada uno de los planos no existe una colimación convencional.
3. **Atenuación** En la detección de fotones por coincidencia, la atenuación dependerá solamente del recorrido total a través del paciente, pero será independiente de la ubicación exacta del evento de aniquilación en la profundidad del tejido. Esto es bastante diferente del caso de SPECT donde la atenuación representa un problema importante debido a la dificultad para corregirla matemáticamente.

La detección de eventos de positrones necesita la llegada al detector de ambos fotones de 511 keV. La pérdida de cualquiera de los fotones debido a la atenuación significa que la detección de coincidencia no se llevará a cabo. Por tanto, vemos que el número de eventos detectados dependerá de la probabilidad de que ambos fotones alcancen los detectores.

La corrección de atenuación en PET se basa en que, independientemente de la localización del evento de aniquilación, uno u otro de los fotones atravesarán la totalidad del espesor corporal. De hecho, lo mismo se aplica para una fuente de positrones colocada fuera del cuerpo, situación en la cual uno de los fotones no será atenuado mientras que el otro deberá atravesar la totalidad del cuerpo. Este hecho permite una medida directa de la atenuación para cada trayecto del rayo gamma que atraviesa el cuerpo utilizando una fuente de transmisión externa.

4. Detección

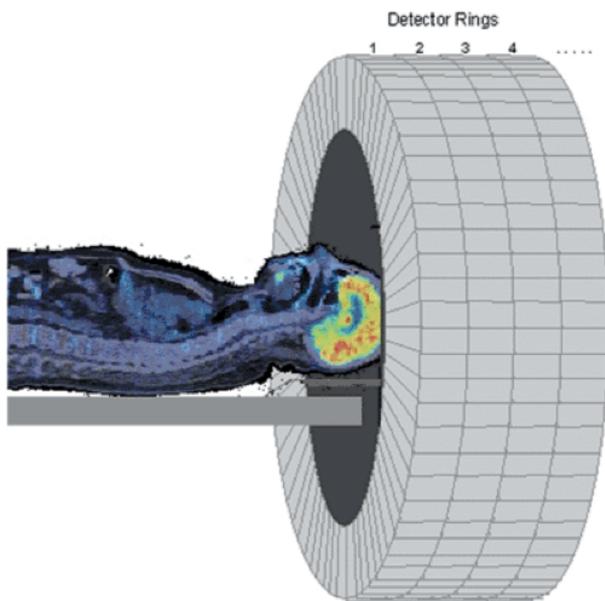


Figura 3.6: Dibujo esquemático de un escáner PET de anillo completo. Un campo de visión cilíndrico está encerrado por anillos de detectores que registran los fotones de aniquilación. La extensión axial cubierta por los anillos detectores es típicamente 15 a 25cm

Los detectores se utilizan para medir la cantidad de energía de los fotones. Los más comunes hoy en día se basan en una combinación de fotodetector y cristal centelleador. Los cristales centelleadores detectan los fotones de 511 keV de alta energía resultantes de la aniquilación de positrones y, a su vez, generan una cantidad de fotones de centelleo proporcionales a la energía del fotón resultante de la aniquilación. Son fotones mucho más cercanos al rango visible (1 eV).

Los cristales centelleadores junto con los fotodetectores se empaquetan en detectores de bloque que actúan en dos fases, primero detectando los fotones de aniquilación y generando los fotones de centelleo y luego convirtiéndolos en corriente eléctrica. Estos bloques están agrupados círculo y varios de ellos forman un cilindro, el cilindro detector del escáner PET.

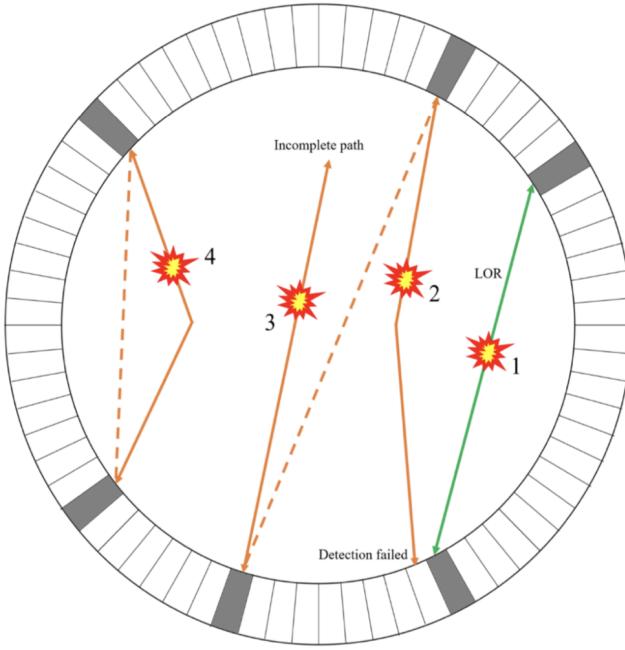


Figura 3.7: Ejemplo de evento coincidente

Para poder crear la imagen, es obligatorio que el anillo detecte ambos fotones de aniquilación. Si no se encuentra ninguno obstáculo, viajarán en direcciones antiparalelas y se detectarán aproximadamente al mismo tiempo.

Esto se llama un evento coincidente, pero varios factores pueden desviar al fotoén de su camino.

Solo se aceptan coincidencias verdaderas, es decir, coincidencias que resultan de la detección correcta de los dos fotones de un mismo evento de aniquilación (línea de respuesta, LOR). El resto de las coincidencias se descartan.

Estas LOR se almacenan como imágenes de proyección llamadas sinogramas. Los sinogramas incluyen los datos sin procesar de la adquisición, y antes de reconstruir la imagen, se realizan correcciones en estos datos. La corrección de atenuación (AC) es necesaria para mitigar los efectos de la dispersión de fotones y las variaciones de eficiencia del detector.

5. **Reconstrucción** Los sinogramas obtenidos en la adquisición contienen información de la imagen, pero es necesario reconstruirla.

El proceso de reconstrucción se realiza mediante algoritmos similares a los utilizados en

TC, por ejemplo FBP o algoritmos iterativos.

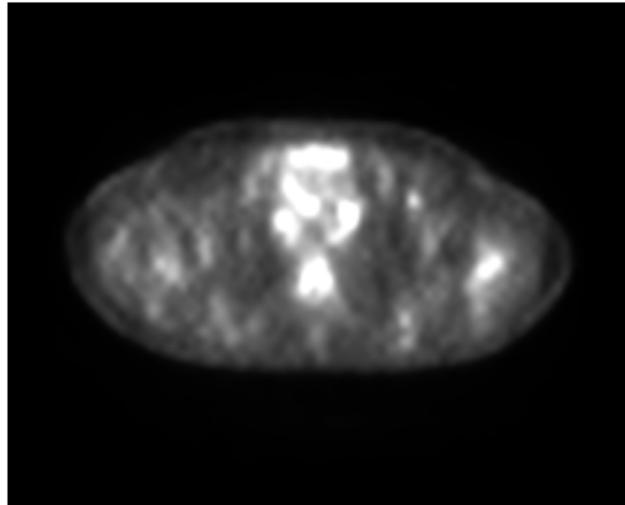


Figura 3.8: Adquisición previa a la reconstrucción

3.1.3.2. Imagen Hibrida y Nuevas Formas de Imagen

Las limitaciones de las técnicas de imagen no invasivas que proporcionan información anatómica (fundamentalmente TC y RM) o funcional (SPECT y PET) han llevado a la idea reciente de combinar en una sola sesión varias modalidades de imagen. La posibilidad de utilizar simultáneamente varias técnicas permite aumentar la escasa resolución espacial de las técnicas nucleares y, por otro lado, aumentar la sensibilidad y la reproducibilidad de la señal de las sondas moleculares detectada por técnicas anatómicas. De esta forma, se optimiza el análisis y la interpretación de las imágenes. Los sistemas híbridos, por lo tanto, permiten aprovechar las ventajas individuales de cada modalidad e integrar la información molecular, anatómica y fisiológica en una sola imagen[16].

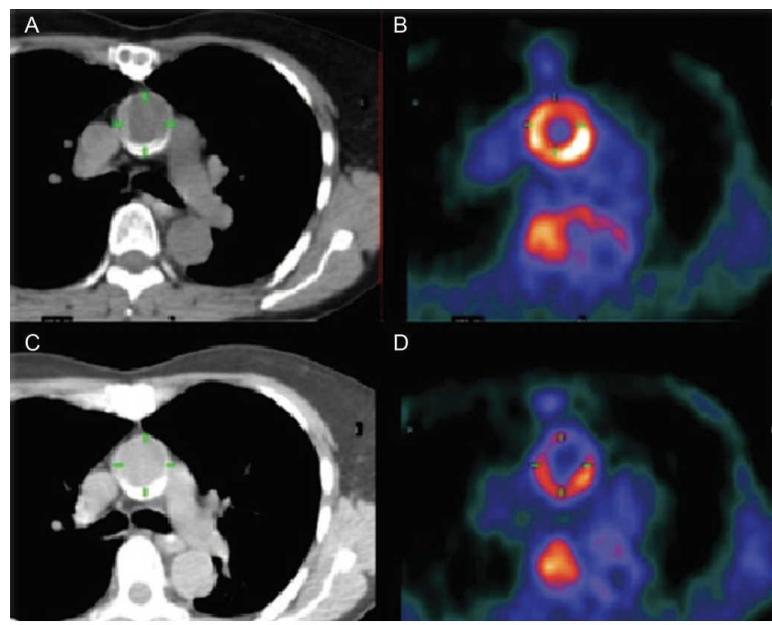
Imagen híbrida: La combinación de dos tecnologías de la imagen, llamada co-registración, imagen de fusión o imagen híbrida, permite visualizar la información proporcionada por dos tipos diferentes de imágenes en un solo set. PET/CT y SPECT/CT, la combinación de PET o SPECT con la TAC (CT), se han convertido en técnicas diagnósticas de uso cotidiano al proporcionar información detallada tanto de la anatomía como de la funcionalidad de órganos y tejidos. Nuevos tipos de imagen híbrida están en uso o en desarrollo, como la PET/RMN, la PET/ecografía, y varias técnicas ópticas combinadas con tecnologías de imagen convencional.

Las modalidades de imagen disponibles actualmente se pueden combinar de diferentes maneras; sin embargo, las plataformas más utilizadas son la combinación de SPECT/TC, PET/TC

y PET/RM (Figura 3.9). La combinación de imagen nuclear con TC se utiliza más frecuentemente en estudios vasculares, por su gran resolución espacial (Figura 3.10) , mientras que la fusión con RM se prefiere en estudios cardiacos que precisan información funcional (función, perfusión y/o realce tardío), pues se evita la radiación de los estudios tomográficos[17].



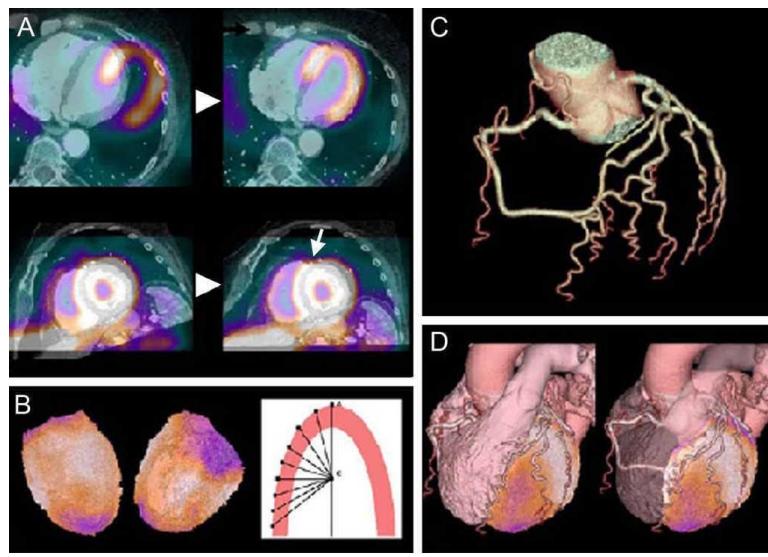
Figura 3.9: Sistema híbrido que integra en una misma sala un equipo de tomografía por emisión de positrones y resonancia magnética compartiendo una mesa giratoria que permite mantener al paciente en la misma posición para tomar las imágenes de tomografía por emisión de positrones y resonancia magnética, con el fin de facilitar la superposición anatómica de las imágenes. PET: tomografía por emisión de positrones; RM: resonancia magnética.



Rev Esp Cardiol. 2013;66:134–43

Figura 3.10: PET/TC de aorta antes (arriba) y durante (abajo) terapia antiaterosclerosis. Obsérvese la reducción en la absorción de FDG en la pared aórtica.

El valor añadido de las técnicas híbridas se debe a su capacidad para lograr en las imágenes fusionadas una correlación espacial de la imagen estructural y funcional, lo cual facilita la interpretación de las imágenes en cuanto a localización de la lesión y su relevancia fisiológica (Figura 3.11). Es imprescindible que haya una superposición perfecta de ambas imágenes, puesto que el mal alineamiento puede conducir a errores en la determinación de las áreas patológicas. También es requisito de estos equipos que corrijan los artefactos de atenuación de los tejidos blandos que afectan a la imagen nuclear mediante las técnicas anatómicas.



Rev Esp Cardiol. 2013;66:134-43

Figura 3.11: PET / TC de aorta antes (arriba) y durante (abajo) terapia antiaterosclerosis. Obsérvese la reducción en la absorción de FDG en la pared aórtica.

Recordando sus Principales Rasgos:

- Combinación de diferentes tecnologías de adquisición de imagen médica: Multimodalidad (Imagen Híbrida).
- Combinación de diferentes energías en las ondas electromagnéticas utilizadas para generar la imagen: Multiespectral.

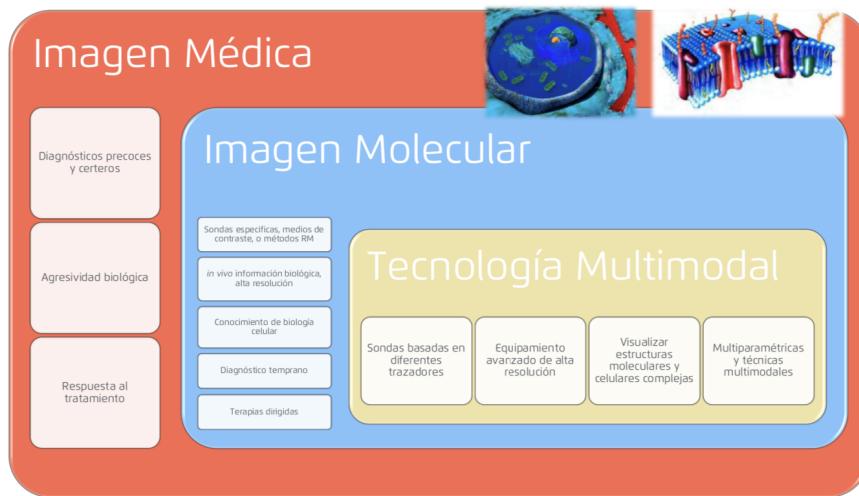


Figura 3.12: Rasgos principales de imagen médica

Tipos de Imagen Híbrida

- PET/TC
- PET/RM

PET/TC

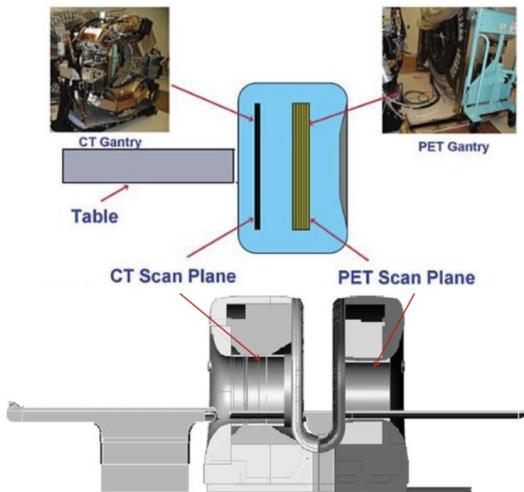


Figura 3.13: PET/TC

La **PET/TC** proporciona una combinación de la información metabólica de las imágenes PET y la información morfológica de la imagen TC en una simple exploración. Como se ha demostrado en múltiples situaciones clínicas, la información obtenida en los estudios PET/TC es de mayor exactitud en la evaluación de los pacientes oncológicos que las imágenes PET o TC por separado, aun cuando se interpreten de forma conjunta[18].

Las imágenes del cuerpo humano obtenidas con esta técnica se utilizan para evaluar diversas enfermedades, entre las que cabe destacar:

- Detección de cáncer y estudio del efecto del tratamiento mediante la caracterización bioquímica de los cambios en el tejido canceroso.
- Determinación del flujo sanguíneo al corazón y evaluación de los signos de enfermedad coronaria. Se usa también para determinar cuáles zonas de funcionamiento reducido están aún vivas, y cuáles son tejido cicatrizal causado por un infarto previo. Junto con el estudio de perfusión miocárdica, los PET permiten diferenciar entre el músculo cardíaco no funcional y el músculo cardíaco todavía viable que se podría beneficiar de un procedimiento como angioplastia o bypass (puente) coronario, que restablecen el flujo sanguíneo y mejoran el funcionamiento del corazón.
- Evaluación de los pacientes con trastornos de la memoria de causa no determinada, tumores de cerebro sospechados o comprobados, o convulsiones que no responden al tratamiento médico y por lo tanto son candidatas para cirugía.

PET/TC en Oncología

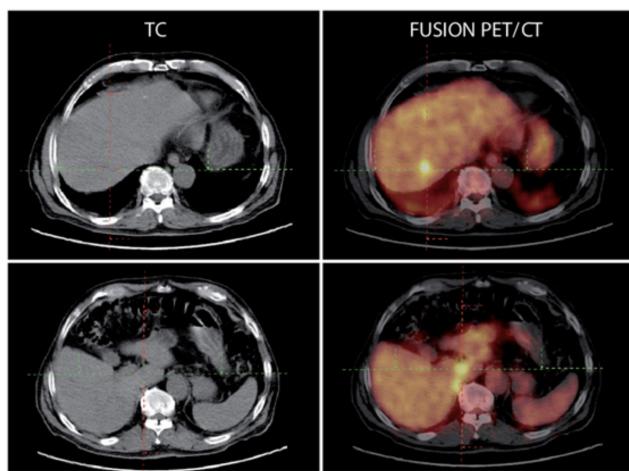


Figura 3.14: PET/CT abdominal con depósitos hipermetabólicos en lóbulo hepático derecho de moderada intensidad (SUV máx segmento VIII: 5.7, segmento VI: 3.62, lóbulo caudado: 5.28, y cúpula hepática: 3.16).

Objetivo:

El objetivo diagnóstico es visualizar el aumento del metabolismo de glucosa en órganos o tejidos concretos en las siguientes indicaciones:

- Diagnóstico.
- Estadificación.
- Monitorización de respuesta a tratamiento.
- Detección en caso de sospecha razonable de recidiva.

Indicaciones Claramente Establecidas:

- Detección de tumores malignos, dadas las diferencias fisiopatológicas y en la actividad biológica que existen entre procesos benignos y malignos.
- Estadificación y re-estadificación, dada la posibilidad de realizar estudios de cuerpo entero en una misma exploración.
- Localización de tumor primario desconocido.
- Detección de recurrencias tumorales.
- Diferenciación entre recidiva y radio-necrosis.
- Detección de segundo tumor primario (especialmente en cánceres de cabeza y cuello).
- Predicción de la respuesta a la quimioterapia.
- Monitorización del tratamiento. Permite modificar precozmente el tratamiento en aquellos casos con escasa o nula respuesta.

PET/TC en Cardiología

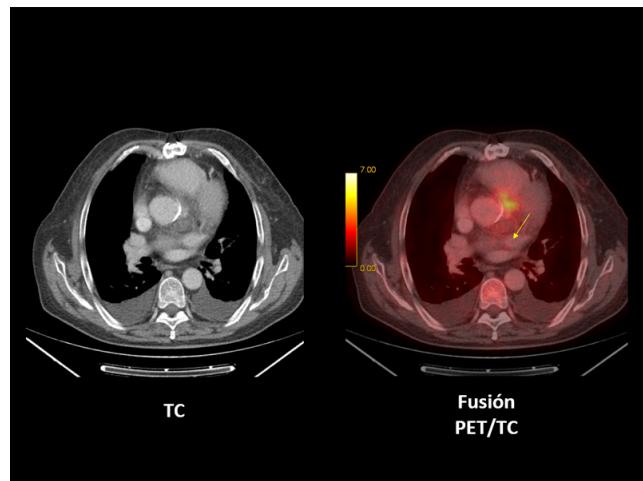


Figura 3.15: PET/TC en Endocarditis

Objetivo:

El objetivo diagnóstico es visualizar el tejido miocárdico viable para evaluar la viabilidad miocárdica en pacientes con disfunción grave del ventrículo izquierdo y que son candidatos a revascularización, sólo cuando las técnicas de imagen convencionales de flujo sanguíneo no son concluyentes.

El SPECT con Tl 201 o compuestos con Tc99m (Sestamibi o Tetrofosmin) es una herramienta muy popular para evaluar la perfusión miocárdica. A pesar de ser un método eficaz, comparado con PET, SPECT cardiológico tiene algunas limitantes. Una de las limitantes de PET en cardiología es la disponibilidad de los radiotrazadores (Rb 82, o N13).

Aplicaciones Clínicas:

- Evaluación de viabilidad miocárdica
- Perfusión miocárdica: en pacientes cuyas gammagrafías de Tc99m Sestamibi fueron inconclusas
- Potencial evaluación de la Enfermedad Arterial Coronaria en pacientes sintomáticos y asintomáticos

PET/RM

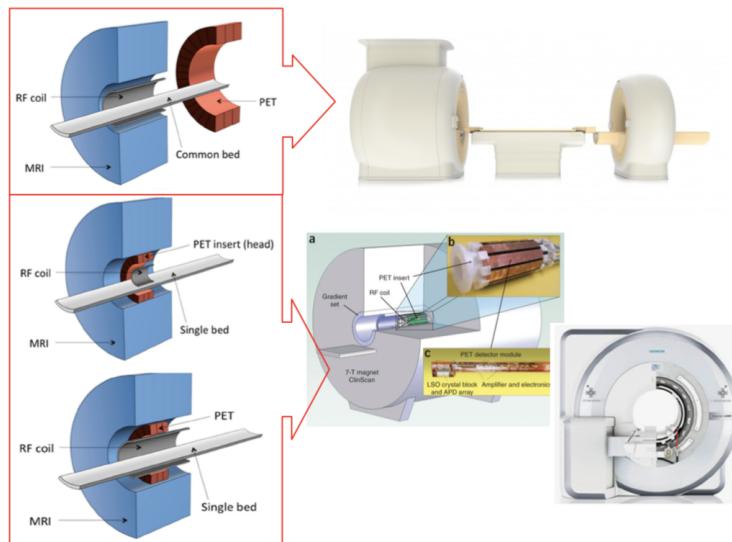


Figura 3.16: PET/RM

La **PET/RM** representa un modelo singular de Diagnóstico por la Imagen. Aúna, en una única exploración, la información diagnóstica de la imagen de los procesos bioquímicos y moleculares que ofrece la Tomografía de Emisión de Positrones (PET) con la caracterización morfológica y morfo-funcional tisular que ofrece la Resonancia Magnética (RM)[19].

La PET/RM permite identificar la distribución corporal y sectorial de diferentes trazadores metabólicos (F18-FDG, F18-Fluorocolina, F18-Fluoruro, F18-Fluorodopa, Ga68-DOTA...). Al mismo tiempo, correlaciona simultáneamente la distribución de estos radiofármacos con la caracterización estructural y funcional que proporcionan las secuencias de un equipo RM 3T.

PET/RM en Oncología

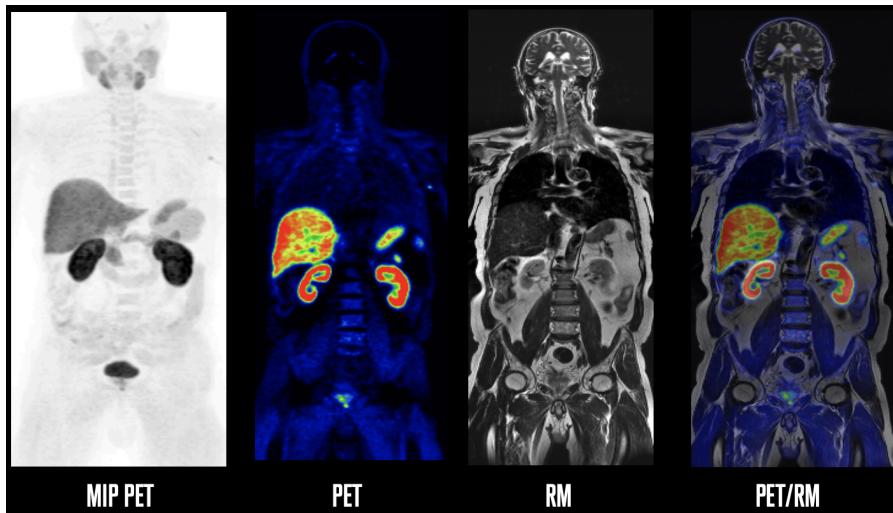


Figura 3.17: Las innovaciones en las técnicas de imagenología PET-RM están demostrando ser útiles para el diagnóstico no invasivo de condiciones médicas, tales como la demencia temprana (Fotografía por Getty Images).

Principales Elementos de Valor Diferencial:

- Rapidez y comodidad para el paciente
- Evaluación integral y personalizada
- Diagnóstico más preciso
- Posibilidad de uso de trazadores específicos
- Menor radiación
- Simplificación del algoritmo diagnóstico

Indicaciones claramente establecidas:

En todas aquellas patologías oncológicas en las que esté indicada la realización de una PET y en las que la RM sea una herramienta diagnóstica superior a la TC, la PET/RM es una herramienta diagnóstica superior a la PET/TC.

Destacamos:

- Cáncer de próstata
- Cáncer ginecológico
- Patología tumoral hepática
- Oncoendocrinología
- Cáncer de mama
- Oncología cerebral

PET/RM en Neurología:

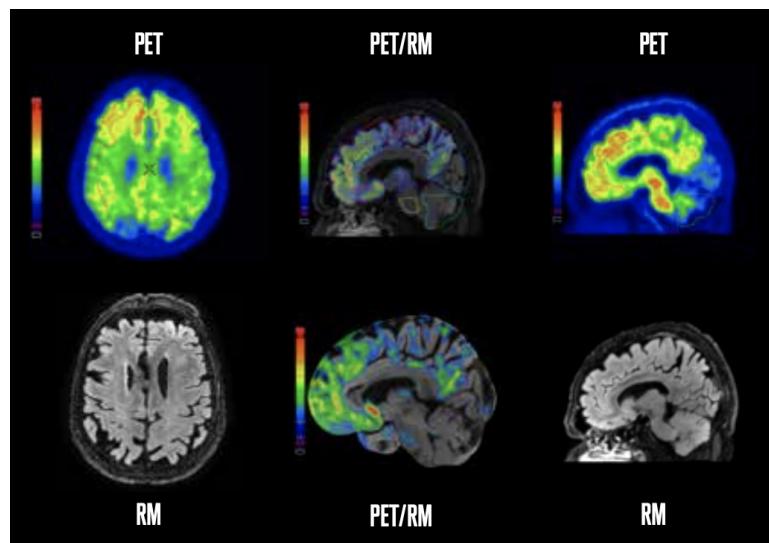


Figura 3.18: PET/RM en neurología

El concepto de enfermedad de Alzheimer ha cambiado en la última década. Esta enfermedad se inicia en una fase silente o asintomática, pasa por una fase prodrómica de deterioro cognitivo leve y finaliza en la fase de demencia. La RM determina la atrofia hipocampal en el inicio del proceso neurodegenerativo. La tomografía cerebral PET con fluorodesoxiglucosa (FDG) permite ver, en la misma fase prodrómica, hipometabolismo en determinadas áreas del cerebro. Ambas técnicas son biomarcadores del inicio de la enfermedad y permiten valorar objetivamente la severidad de la neurodegeneración.

Un biomarcador todavía más precoz es la PET cerebral con trazadores marcadores de β amiloide. Permite detectar el depósito de β amiloide cerebral en todas las fases de la enfermedad,

incluso en la asintomática. Hay que destacar que la negatividad de un estudio PET cerebral con trazadores marcadores de β amiloide excluye, con precisión muy elevada, el diagnóstico de enfermedad de Alzheimer. La PET/RM con trazadores marcadores de proteína Tau fosforilada, aún sin uso clínico, **aumentará todavía más la precisión diagnóstica** de la enfermedad de Alzheimer.

El paradigma descrito para PET/RM en enfermedad de Alzheimer puede ser extrapolado a otras enfermedades neurodegenerativas. También la patología epiléptica y el estudio de tumores cerebrales tienen lugar en la PET/RM. Puede emplearse la PET/RM con F18- fluorodopa para el diagnóstico diferencial de la enfermedad de Parkinson, sobre todo en su fase inicial y para hacer diagnóstico diferencial con los parkinsonismos. En tumores cerebrales ya tratados, de evolución complicada, en los que se plantean los diagnósticos de recidiva vs pseudoprogresión, la PET/RM puede ser útil empleando trazadores oncofílicos, como la F18-fluorocolina; F18-fluorodopa; F18-fluorotimidina.

La PET/RM es una herramienta cargada de futuro, ya que podrá utilizar nuevos y prometedores radiofármacos, marcadores de alteraciones proteonómicas y genómicas celulares específicas de las distintas enfermedades del sistema nervioso central.

PET/RM en Cardiología:

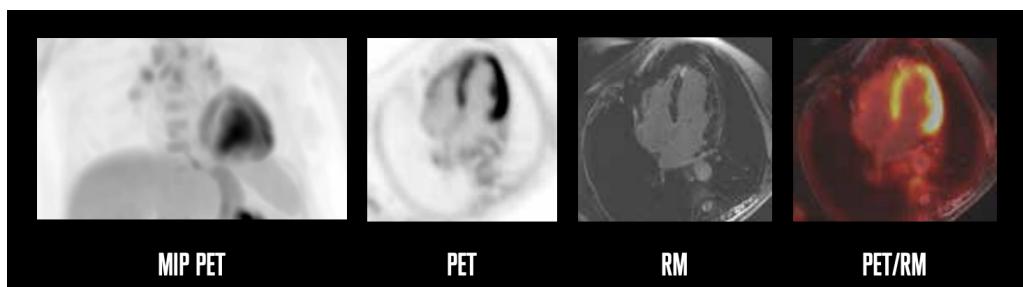


Figura 3.19: PET/RM en Neurología

La cardiorresonancia magnética (CRM) es, sin duda, una herramienta fundamental en el diagnóstico de patología cardiovascular. Se considera actualmente el patrón oro para el estudio de la morfología y función ventricular, y proporciona información de caracterización tisular para detección de edema, necrosis y fibrosis focal. Asimismo, progresivamente aumenta su utilización en estudios de perfusión miocárdica, con buen rendimiento diagnóstico.

La PET, igual que la SPECT cardiaca, es una técnica de alto rendimiento en la patología isquémica coronaria y en la identificación del o de los “vasos culpables”. Se reconoce además el valor complementario de la PET, respecto a la CRM, en la determinación de la viabilidad

miocárdica. La PET es una herramienta diagnóstica de creciente interés en el estudio de la patología miocárdica inflamatoria (miocarditis, sarcoidosis...), de la infección (endocarditis sobre válvula nativa) y del diagnóstico de determinadas enfermedades cardíacas de depósito como la amiloidosis.

La fusión de estas dos técnicas, PET y CRM, puede, por tanto, **proporcionar un estudio cardiovascular integral y una mejor caracterización de determinadas enfermedades**, incluyendo el campo de la Cardio-oncología.

Perspectivas de Futuro en la Imagen Híbrida

Las perspectivas de futuro nos llevan a los sistemas híbridos que combinan la PET y la TCAR. La fusión PET + TCAR ó PET + RM se realizaba hasta ahora de manera retrospectiva mediante un sistema de corregistro por software. La fusión de imágenes PET + TCAR es sumamente útil en oncología, ya que permite comprobar que el depósito patológico de un trazador, como la FDG, coincide con la masa o adenopatía aumentada de tamaño detectada por la TCAR. Esto facilitaría la toma de muestras y la guía de biopsias o, incluso, la planificación de la radioterapia. Los sistemas híbridos permiten someter al paciente a las dos pruebas durante la misma sesión exploratoria.

3.2. Modalidades de Adquisición No Ionizantes:

■ Técnicas No Ionizantes:

- Resonancia Magnética (RMN)
- Ultrasonido (US)

3.2.1. Resonancia Magnética (RMN)

El camino hacia el descubrimiento de la Resonancia Magnética comenzó con las primeras investigaciones sobre la naturaleza del átomo.

En 1897, el físico J. J. Thomson (1856 - 1940), profesor en la Universidad de Cambridge (Inglaterra), descubrió una nueva partícula y demostró que era aproximadamente mil veces más ligera que el hidrógeno. Esta partícula sería bautizada con el nombre de electrón.

Durante las dos décadas siguientes, una serie de destacados físicos, entre los que destacamos [20,21]:

- Nikola Tesla (1856 - 1943): Físico estadounidense de origen serbio. Trabajó bajo las órdenes de Thomas A. Edison. Entre sus inventos más importantes destaca la bobina de Tesla (transformador capaz de crear un campo de alta tensión y alta frecuencia). Las unidades de inducción de campo magnético empleadas en resonancia magnética llevan su nombre.
- Sir Joseph Larmor (1856 - 1942): Físico y matemático que hizo importantes avances en el campo de la electricidad, la dinámica y la termodinámica así como la teoría electrónica de la materia. Entre sus descubrimientos más importantes destacamos el descubrimiento de la tasa a la que se irradia energía por parte de un electrón acelerado y fué el primero en explicar la separación de líneas espectrales por un campo magnético. Formuló la Ecuación de Larmor, la cual define la frecuencia de precesión del momento magnético nuclear.
- Max Planck (1858 - 1947): Físico alemán mundialmente conocido por inventar las bases de la Teoría Cuántica.
- Ernest Rutherford (1871 - 1937): Físico y químico británico. Investigó sobre la detección de las radiaciones electromagnéticas y sobre la ionización del aire producida por los rayos X. Estudió las emisiones radioactivas y logró clasificarlas en rayos alfa, beta y gamma.

En 1902, en colaboración con Frederick Soddy, Rutherford formuló la teoría sobre la radioactividad natural asociada a las transformaciones espontáneas de los elementos.

En 1911 describió un nuevo modelo atómico (Figura 3:20), que posteriormente sería perfeccionado por Niels Bohr. Según este modelo existía un núcleo central pesado y cargado positivamente y una envoltura o corteza de electrones (carga eléctrica negativa).

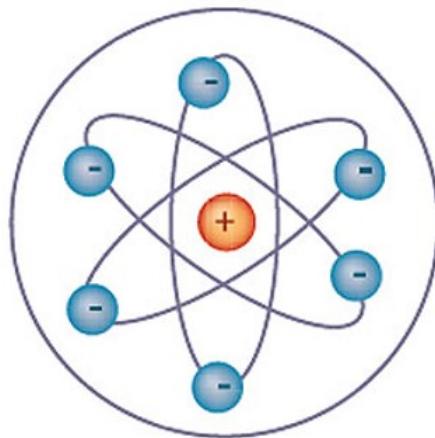


Figura 3.20: Modelo atómico de Rutherford

- Niels Bohr (1885 - 1962): Físico de origen Danés. En el año 1922 fue galardonado con el Premio Nobel de Física por su investigación acerca de la estructura de los átomos y la radiación que emana de ellos.
- Erwin Schrodinger (1887 - 1961): Físico austriaco que compartió el Premio Nobel de Física del año 1933 con Paul Dirac por su contribución al desarrollo de la mecánica cuántica.
- Werner Heisenberg (1901 - 1976): Físico alemán que formuló el principio de incertidumbre o de indeterminación de Heisenberg, una aportación decisiva para el desarrollo de la mecánica cuántica. Revolucionaron la física y elaboraron una teoría y un lenguaje nuevos conocidos como mecánica cuántica.

Gracias a Jean Baptiste Joseph Fourier (1768 - 1830) ingeniero y matemático francés, es posible el procesado de la información de frecuencia y fase de los datos de Resonancia Magnética y utilizar dicha información para la reconstrucción de imágenes.

Fourier realizó experimentos sobre la propagación del calor desarrollando la teoría general de transformación matemática para el análisis de la transferencia de calor entre cuerpos sólidos.

A partir de esta teoría desarrolló la denominada serie de Fourier, de gran importancia en el posterior desarrollo del análisis matemático, y con interesantes aplicaciones a la resolución de numerosos problemas de física.

En el año 1946, dos científicos (Bloch y Purcell) descubrieron el fenómeno de la Resonancia Magnética Nuclear de manera simultánea e independiente.

No fueron los únicos científicos que trabajaban en este campo.

En 1924 Wolfgang Pauli (1900 - 1958) sugirió la posibilidad de la existencia de un espín nuclear intrínseco. Al año siguiente George Eugene Uhlenbeck (1900 - 1988) y Samuel A. Goudsmit (1902 - 1978) introdujeron el concepto del electrón giratorio.

Dos años más tarde, Pauli y Charles Galton Darwin (1887 - 1962) desarrollaron un marco teórico para encajar el concepto de espín electrónico en la nueva mecánica cuántica desarrollada por Erwin Schrodinger y Werner Heisenberg.

Después de unos trabajos innovadores, en 1933, Otto Stern (1888 - 1969) junto con Walther Gerlach (1889 - 1979) fueron capaces de medir el efecto del espín nuclear por desviación de un haz de moléculas de hidrógeno (efecto Stern - Gerlach).

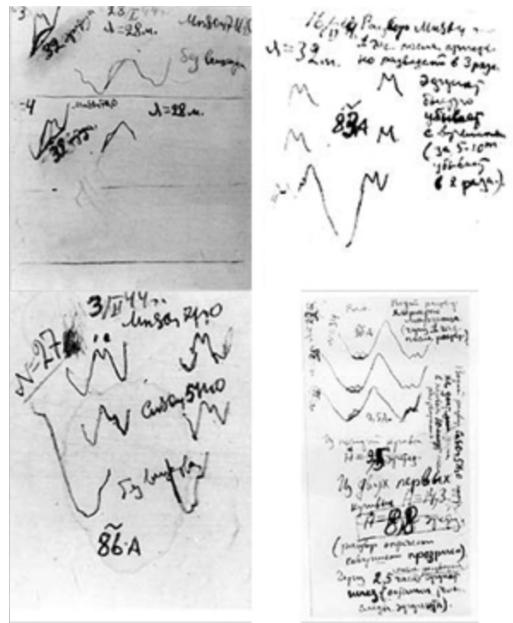


Figura 3.21: Primeras señales de REE

Por otro lado, en 1944 Yevgeni K. Zavoisky (1907 - 1976) pudo registrar señales de la resonancia del espín electrónico (Figura 3.21).

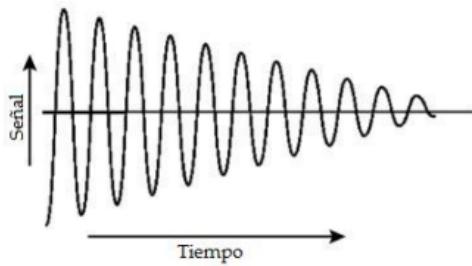


Figura 3.22: Free induction decay (FID)

En 1949 E. Hahn demostró que era posible la obtención de una señal (FID: Free Induction Decay, Figura 3.22) utilizando pulsos de radiofrecuencia (RF). El método de pulsos de Hahn es la base de la metodología actual en Resonancia Magnética.

En 1950 Proctor y Yu descubrieron el efecto de desplazamiento químico (núcleos del mismo tipo tienen diferentes frecuencias de resonancia, Figura 3.23).

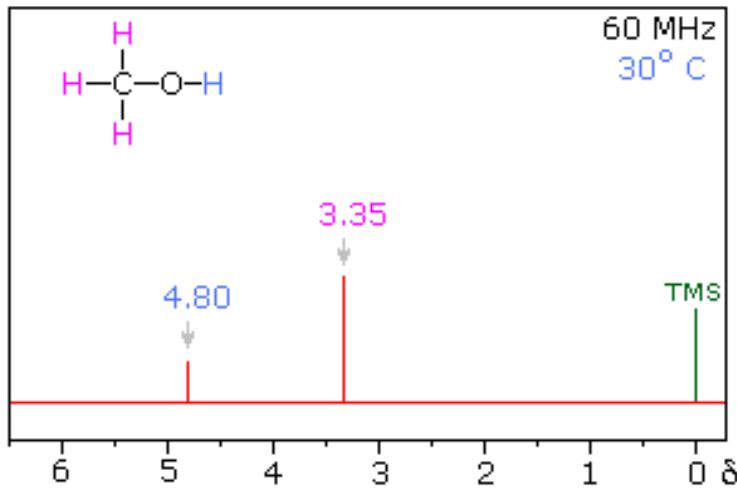


Figura 3.23: Desplazamiento químico de H

En 1971 Raymon V. Damadian publicó en Science el trabajo titulado “Tumor detection by nuclear magnetic resonance”. En este proyecto se propone utilizar las medidas resonancia spin-echo como un método para discriminar entre tumores malignos y tejido normal.

Damadian encontró diferencias en T1 (tiempo de relajación longitudinal: tiempo que tarda la magnetización longitudinal en recuperarse) y T2 (El T2 de un tejido es el tiempo que emplea en perder el 63 % de su magnetización transversal) entre seis muestras de tejido normales y dos tumores sólidos del hígado y el riñón de una rata.

Además, comparó fibroadeno mas con dichos tumores, encontrando que también se diferenciaban de ellos. Así, los valores de T1 y T2 de tumores se encontraban fuera del rango de los valores normales y de los fibroadenomas. Damadian apodó al método analítico para diferenciar los valores de relajación “FONAR” (Field Focused Nuclear Magnetic Resonance).

Sus resultados los corroboró con tejido humano en 1974. En 1976 junto a su equipo, construyó el primer tomógrafo de Resonancia Magnética de cuerpo entero al que llamaron ‘El Indomable’ (Figura 3.24), obteniendo la imagen del tumor de una rata publicada en la revista Science.



Figura 3.24: Primer tomógrafo de RM: "El Indomable"

El avance fundamental para la medicina se produjo cuando el químico estadounidense Paul Lauterbur (1929 - 2007), en la Universidad de New York, tuvo la idea de codificar espacialmente la señal mediante la aplicación de gradientes magnéticos y después reconstruir la imagen en forma similar a la tomografía computada.

En su trabajo “Image formation by induced local interactions; example employing magnetic resonance”, publicado en la revista Nature en 1973, mostró que agregando campos magnéticos adicionales al campo principal y obteniendo un conjunto de proyecciones de la distribución de la señal de dos tubos de prueba conteniendo agua normal dentro de un contenedor de agua deuterada, se podía reconstruir una imagen por medio de retroproyección filtrada (similar a la tomografía computada), Figura 3.25. Un año más tarde publicó la imagen de una cavidad torácica de un ratón.

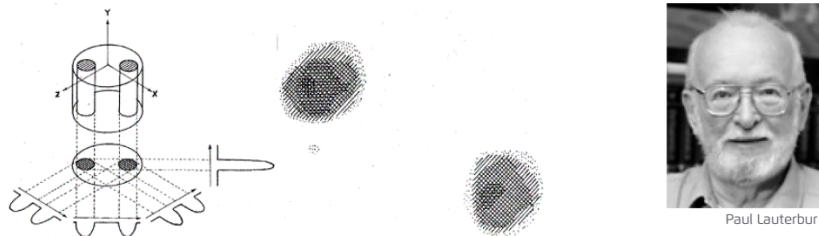


Figura 3.25: Primera imagen de RM de dos tubos llenos de agua

Peter Mansfield (Figura 3.26), físico inglés, en 1973 descubrió que el uso de gradientes de campo magnético producía señales que podrían ser analizadas directamente para proveer la información espacial, incluyendo una descripción matemática de la transformación de una señal temporal a la representación espacial, introduciendo el concepto del espacio-K. Posteriormente siguió trabajando en la selección del corte y en 1976 propuso las secuencias EPI (eco-planar imaging), o como llenar el espacio-K en forma rápida, en un solo disparo (“single shot”).



Figura 3.26: Peter Mansfield

Richard Ernst aplicó la transformada de Fourier bidimensional a la recepción de señales espectroscópicas y demostró que utilizando gradientes de campo junto con la transformada de Fourier bidimensional era posible reconstruir una imagen.

En 1981 se instaló en Londres el primer prototipo de tomógrafo por RM nuclear para uso hospitalario en el Hammersmith Hospital.

En España el primer equipo de RM se instaló en el año 1983 en el Centro Médico de Resonancia Magnética de Barcelona, obteniéndose las primeras imágenes (Figura 3.27) el 14 de Diciembre de 1983 bajo la dirección de los Dres. Jaume Gili y Manuel Sanz.



Figura 3.27: Primera imagen de RM en España

Fundamentos de la Resonancia Magnética

Los estudios con imágenes por resonancia magnética (IRM) usan un gran imán y ondas de radio para observar órganos y estructuras que se encuentran en el interior del cuerpo.

Estas imágenes se utilizan con la finalidad de diagnosticar una variedad de afecciones, desde rupturas de ligamentos hasta tumores.

Durante el escaneo, el paciente se acuesta en una mesa que se desliza dentro de un aparato con forma de túnel. La realización del estudio suele tardar un determinado tiempo y el paciente debe permanecer inmóvil y dicho estudio es indoloro para el paciente.

A continuación se indica una imagen perteneciente a un equipo de resonancia magnética (Figura 3.28).



Figura 3.28: Equipo de Resonancia Magnética

Para entender el fundamento de la resonancia magnética [22] nuclear es necesario el estudio de:

- **Espín nuclear:** Los tres tipos de partículas que constituyen el átomo (electrones, neutrones y protones) tienen una propiedad intrínseca llamada espín: Los electrones del átomo circulan alrededor del núcleo y generan un campo magnético que llevará asociado un determinado momento angular. Existe un momento angular asociado a cada partícula y éste se describe mediante el número cuántico de espín que puede tomar valores de $+\frac{1}{2}$ o $-\frac{1}{2}$.

Para la Resonancia Magnética Nuclear (RMN) es muy importante el espín correspondiente a los protones y neutrones del núcleo atómico.

En el núcleo atómico, cada protón se pueden aparear con otro protón con espín antiparalelo (algo análogo a lo que sucede con los electrones en los enlaces químicos).

Los pares de partículas que resultan de combinar un espín de signo positivo con otro negativo, da como resultado un espín neto de valor cero. Por esta razón núcleos con número de protones y neutrones impar dan lugar a un espín neto, donde el número de desapareamientos contribuye con $\frac{1}{2}$ al total del número cuántico de espín nuclear, denominado I. Por tanto, entre los elementos de la Tabla Periódica, cada isótopo de un determinado

átomo, dependiendo de cual sea el número de protones y neutrones del núcleo, va a tener un determinado valor de I (Figura 3.29).

Num. protones	Num. neutrones	protones + neutrones	spin I	ejemplos
par	par	par	0	12C, 16O
par	ímpar	ímpar	1/2	1H, 13C, 15N
ímpar	par	ímpar	>=1	2D, 14N

Figura 3.29: Table ejemplo num protones y neutrones

Cuando I no es nulo, el núcleo tiene un momento angular de espín y un momento magnético asociado, μ , que depende de la dirección del espín. En los experimentos de RMN modernos lo que se hace es manipular el momento magnético.

El momento de espín angular que un núcleo puede tomar va desde $+I$ a $-I$ en pasos enteros. Este valor se conoce como el número cuántico magnético, m . Para un núcleo dado, el número total de estados posibles del momento angular es $(2I+1)$. El momento angular de espín es una magnitud vectorial. La componente z del mismo (denominada Iz) está cuantizada:

$$Iz = mh/2p$$

$$m = (+I, I - 1, I - 2, \dots, -I)$$

Donde h es la constante de Planck.

El comportamiento de los espines en un campo magnético estático puede ser:

1. Situación de Equilibrio

Una muestra de isótopos de espín $I=\frac{1}{2}$ (por ej 1H, 13C o 19F) tienen dos posibles estados del momento magnético $+\frac{1}{2}(\alpha)$ o $-\frac{1}{2}(\beta)$

En ausencia de campo magnético externo (B_0) los espines se encuentran desordenados pudiendo apuntar en cualquier dirección.

Cuando la muestra es sometida a un campo magnético externo, los distintos núcleos tenderán a orientarse originando dos posibles estados (Figura 3.30, 3.31):

- Alineado con el campo externo (estado α). Es el estado más estable o de menor energía.
- No alineado con el campo externo (estado β). Es el estado menos estable o de mayor energía.

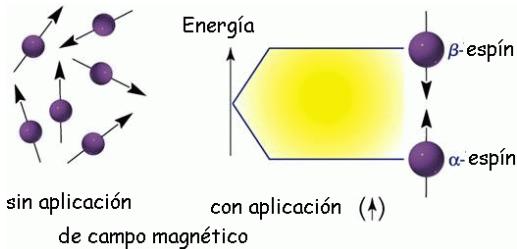


Figura 3.30: Diferencia de energía para los dos estados

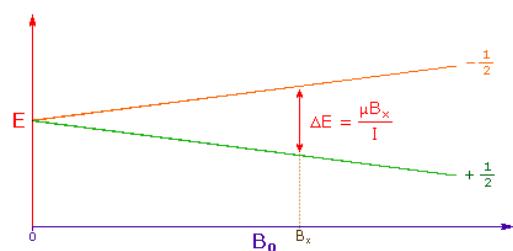


Figura 3.31: Representación gráfica

Los espines nucleares de la muestra tienden a orientarse preferiblemente aunque no exclusivamente en la dirección del campo magnético externo, generándose un pequeño exceso de población en el nivel de menor energía.

La diferencia de energía (ΔE) entre los estados α y β viene dada por la siguiente ecuación:

$$\Delta E = (hyB_0)/2\pi$$

La situación de resonancia entre los dos estados se consigue aplicando una radiación electromagnética (generalmente en la región de las radiofrecuencias, MHz) que tenga exactamente el valor de energía ΔE .

La energía de un fotón es $E = h\nu$, donde ν es su frecuencia.

Por tanto, la frecuencia de la radiación electromagnética requerida para producir resonancia de un determinado núcleo en un campo magnético B_0 viene dada por la frecuencia de Larmor:

$$\nu = yB_0/2\pi$$

2. Situación Fuera de Equilibrio

Un pequeño pulso en la región de las radiofrecuencias (MHz) aplicado en un plano perpendicular al campo magnético del imán genera un segundo campo magnético (dirección B_1) que puede inducir transiciones (cambios de población) entre los estados del espín. Esto sucederá cuando el pulso tenga exactamente la energía exacta ΔE que separa los dos estados $+\frac{1}{2}$ y $-\frac{1}{2}$, o dicho de otro modo, cuando su frecuencia coincida con la frecuencia de Larmor (Figura 3.32).

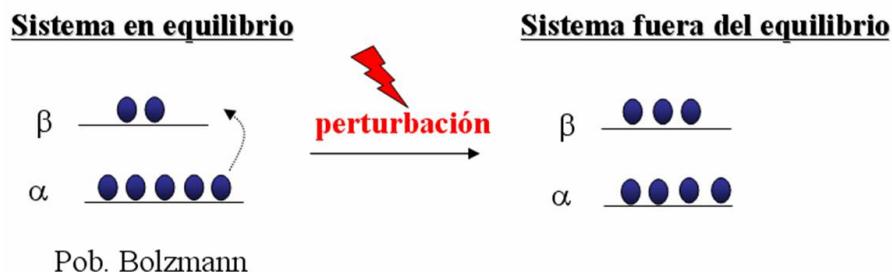


Figura 3.32: Perturbación del sistema en equilibrio

A nivel macroscópico, durante el tiempo que dura el pulso, este segundo campo magnético se producen rotaciones de los espines nucleares "ímanes". Cuando el pulso deja de aplicarse los espines dejan de rotar y quedan alineados (un pequeño exceso de población), en una dirección en principio arbitraria que puede no ser la misma que la situación de equilibrio (Figura 3.33).

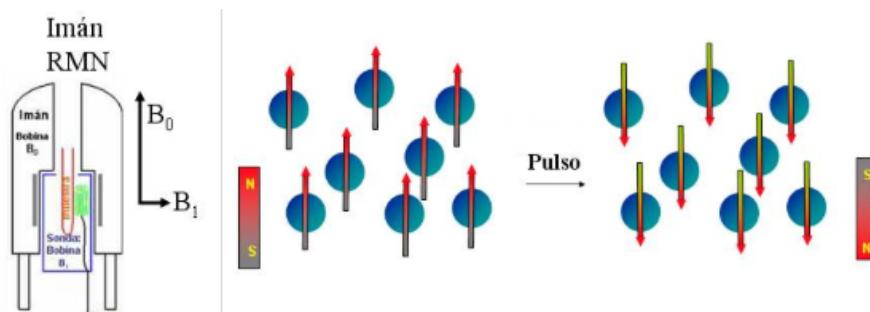


Figura 3.33: Situación creada después de aplicar un pulso

Cuando el pulso cesa, los espines nucleares que se encuentren en situación fuera del equilibrio, tienden a recuperar espontáneamente el estado inicial de población de

equilibrio (donde los distintos núcleos están alineados con el campo externo) emitiendo el exceso de energía en forma de onda de radio (Figura 3.34). Esta señal es amplificada y digitalizada convenientemente y es lo que se conoce como un **espectro de RMN**.

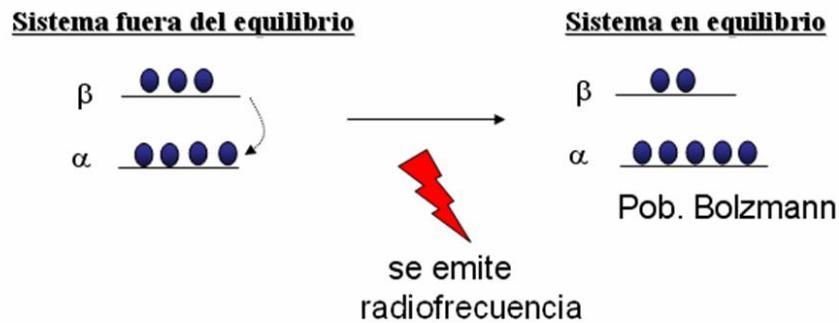


Figura 3.34: Cese del pulso

Apantallamiento Nuclear

El campo magnético que se siente en las inmediaciones del núcleo no es el mismo para todos los núcleos de la muestra. Existen efectos locales de los demás núcleos vecinos, especialmente de aquellos núcleos con espines activos a la RMN, así como un efecto de apantallamiento de campo debido a la nube de electrones circundantes que hacen que el campo efectivo (B_{eff}) que siente cada núcleo sea ligeramente diferente según su entorno químico y menor al del propio campo magnético externo B_0 .

Dado que el campo magnético efectivo en general va a ser diferente para cada tipo de núcleo no equivalente (con entorno químico diferente), también lo va a ser el gap de energía entre sus estados a y b, y en definitiva también su frecuencia de resonancia característica. Las diferencias de apantallamiento en las inmediaciones de cada núcleo va a dar lugar a diferentes picos en el espectro (frecuencias) de RMN.

El desplazamiento Químico (δ) mide la diferencia de frecuencia respecto a un compuesto de referencia. El compuesto de referencia para RMN de 1H y ^{13}C es el Tetrametilsilano, TMS (Si $(CH_3)_4$) que define el cero en la escala de partes por millón (ppm).

Bases Físicas de la RMN

Momento Magnético del Espín Nuclear (μ)

El espín es una componente del momento angular de los núcleos atómicos, electrones (y otras partículas elementales) que no puede ser descrita como parte de su momento orbital.

En física de partículas el momento magnético (μ) asociado con el momento angular orbital (L) de una partícula cargada, por ej. un núcleo atómico, viene dado por Figura 3.35.

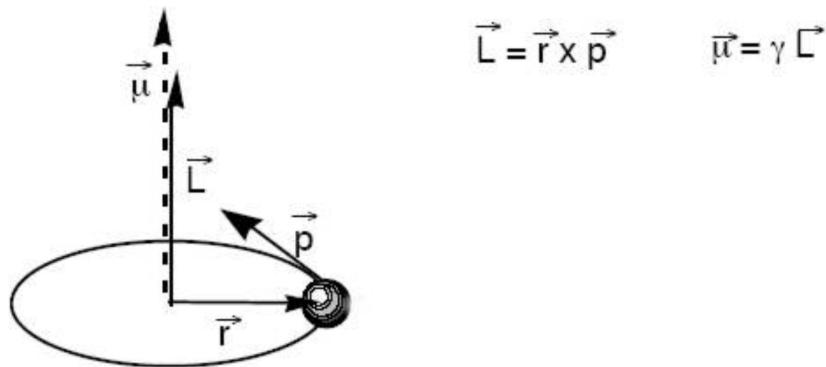


Figura 3.35: Momento magnético asociado con el angular orbital

El momento angular de espín (J) también está asociado con el momento magnético (μ) mediante: $\mu = \gamma J$ siendo γ , es la constante magnetogírica, una propiedad fundamental de cada isótopo nuclear con espín no nulo.

Interacción entre el Momento Magnético del Espín Nuclear y el Campo Magnético Externo B_0

Esta interacción se describe mediante un torque (T) que actúa sobre el momento magnético μ . Como consecuencia el momento magnético μ experimenta un cambio de dirección en el espacio en torno a la dirección del campo magnético B_0 (similar al movimiento de una peonza, Figura 3.36).

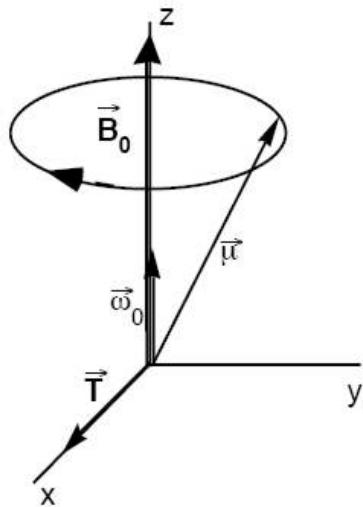


Figura 3.36: Cambio de dirección del momento magnético

La descripción cuántica del efecto tiene como resultado que la componente Z del momento angular de espín (parelola al campo magnético B_0 aplicado) está cuantizada. solo puede asumir ciertos valores según sea el número cuántico magnético m_i .

El número total de estados permitidos es $2I + 1$. Para núcleos con $I = \frac{1}{2}$, (por ej. núcleos como ^1H , ^{13}C , o ^{15}N), sólo hay dos estados posibles $m_I = -\frac{1}{2}$ y $m_I = +\frac{1}{2}$. Cuando este tipo de núcleos se inserta en un campo magnético (B_0), se produce un desdoblamiento en dos niveles de energía cuya separación es proporcional a la fortaleza del campo magnético:

$$\Delta E = -\gamma h \left(\frac{1}{2} - \left(-\frac{1}{2} \right) \right) B_0 = -\gamma h B_0$$

En los campos magnéticos habituales de los espectrómetros de RMN disponibles, la diferencia de energía (ΔE) está en la región de las frecuencias de radio, es decir, dentro del espectro electromagnético en la región de MHz (Figura 3.37).

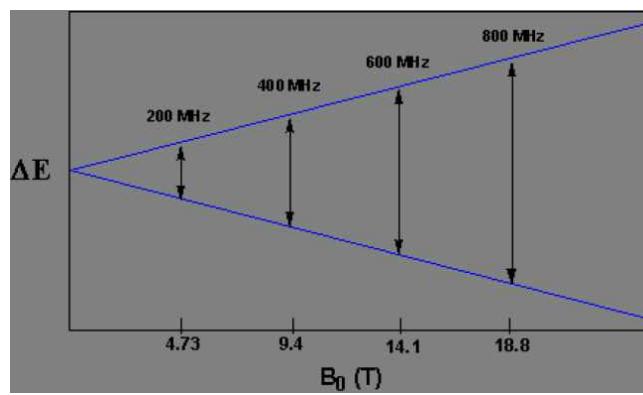


Figura 3.37: Diferencia de energía de los espectrómetros de RMN

Ejemplos de Espectros de RMN (^1H):

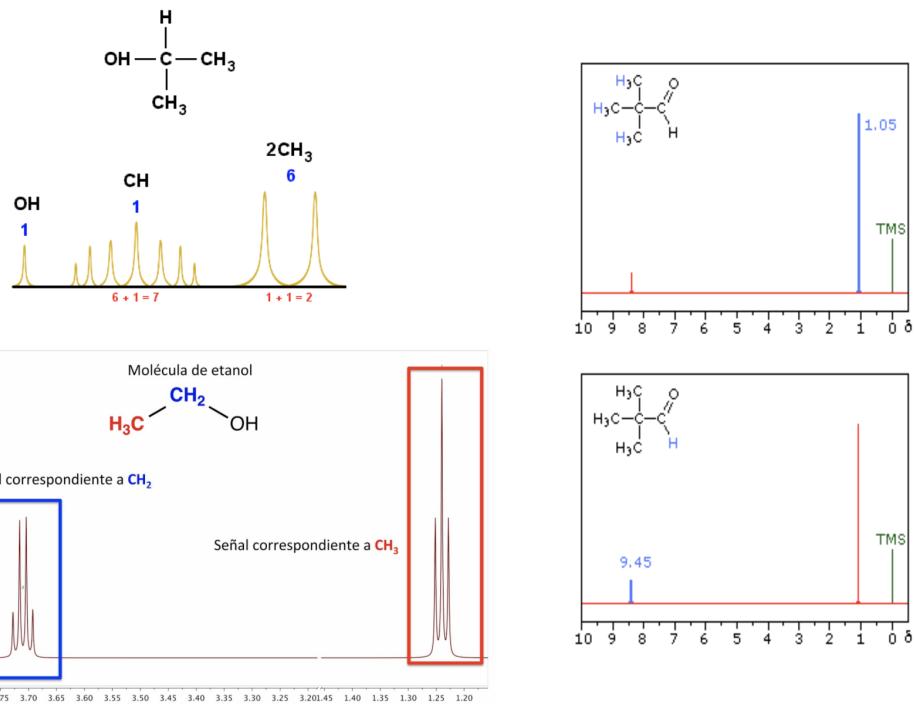


Figura 3.38: Ejemplos de Espectros de RMN

Teniendo en cuenta la teoría explicada anteriormente, podemos concluir que los estudios con imágenes por resonancia magnética (IRM) emplean imanes capaces de generar un campo magnético que obliga a los protones presentes en el cuerpo humano a alinearse con ese campo. Cuando se pulsa una corriente de radiofrecuencia a través del paciente, los protones se estimulan y giran fuera de equilibrio. Cuando cesa esa corriente, los protones vuelven al estado inicial liberando energía que es detectada por los sensores de IRM.

El tiempo que tardan los protones en alinearse con el campo magnético externo, así como la cantidad de energía liberada, cambian dependiendo del entorno y la naturaleza química de las moléculas.

Usos de IRM

Los escáneres de IRM son particularmente apropiados para obtener imágenes de las partes no óseas o de los tejidos blandos del cuerpo. Difieren de la tomografía computarizada (TC), en que no usan la radiación ionizante dañina de los rayos X. El cerebro, la médula espinal y los nervios, así como los músculos, ligamentos y tendones se ven mucho más claros con la IRM que con los rayos X y la TC regulares; por esta razón la IRM se utiliza con frecuencia para obtener imágenes de lesiones de rodilla y de hombro.

En el cerebro, la IRM puede diferenciar entre la materia blanca y la materia gris, y también puede usarse para diagnosticar aneurismas y tumores. Debido a que la IRM no utiliza rayos X u otra radiación, es la modalidad de imágenes preferida cuando se requieren imágenes frecuentes para el diagnóstico o la terapia, especialmente en el cerebro. Sin embargo, la IRM es más costosa que las imágenes de rayos X o de escaneo por TC.

Una clase de IRM especializada es la Imagen por Resonancia Magnética funcional (IRMf). Esta se usa para observar las estructuras del cerebro y determinar qué áreas del cerebro se “activan” (consumen más oxígeno) durante varias tareas cognitivas. Se utiliza para fomentar el entendimiento de la organización del cerebro y ofrece un nuevo estándar potencial para evaluar el estado neurológico y el riesgo neuroquirúrgico.

A continuación podemos encontrar algunos ejemplos de Imágenes por Resonancia Magnética (Figura 3.39 IRM del cerebro, 3.40 IRM de Próstata, 3.41 IRM de Tórax, 3.42 IRM de columna):

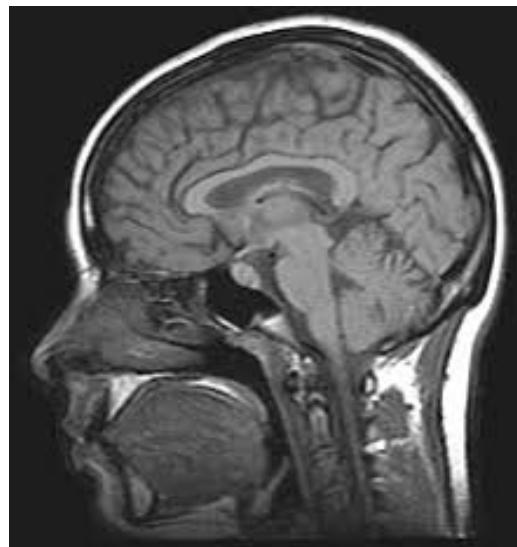


Figura 3.39: IRM del Cerebro

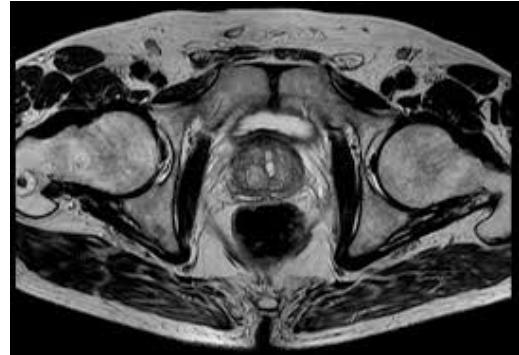


Figura 3.40: IRM de Próstata



Figura 3.41: IRM del Cerebro



Figura 3.42: IRM de Próstata

Riesgos de IRM

Aunque esta técnica no empleo radiación ionizante dañina que se encuentra en otro tipo de técnicas como rayos X y TC si emplea un potente campo magnético. Este campo magnético se extiende más allá de la máquina y ejerce fuerzas sobre objetos de hierro, algunos aceros y otros objetos magnetizables. Por ese motivo, Los pacientes deben notificar a sus médicos de cualquier condición médica o implante que tengan antes de un escaneo por IRM.

Antes de someterse a una IRM, se debe considerar lo siguiente:

- Personas con implantes, en particular aquellos que contengan hierro (marcapasos, monitores cardiacos subcutáneos, bombas de insulina,estimuladores cerebrales profundos y cápsulas endoscópicas no deben someterse a esta técnica.
- Debido al ruido que producen los equipos de resonancia, se requiere protección especial para los oídos.
- En ocasiones se produce una sensación de espasmos debido a los cambios bruscos de campos magnéticos.
- Medios de contraste: los pacientes con insuficiencia renal severa que requieren diálisis podrían correr el riesgo de contraer una enfermedad rara pero seria llamada fibrosis nefrogénica sistémica, que puede estar ligada al uso de ciertos medios que contienen gadolinio, como la gadodiamide y otros. Aunque no se ha establecido una correlación causal, las normas actuales en los Estados Unidos recomiendan que los pacientes de diálisis solo deben recibir medios de contraste con gadolinio cuando sea esencial, y que la diálisis se debe realizar tan pronto como sea posible después del estudio para retirar el medio de contraste del cuerpo con prontitud.
- Embarazo: aunque no se han demostrado efectos en el feto, se recomienda evitar los escaneos de IRM como precaución, especialmente en el primer trimestre del embarazo cuando los órganos del feto se están formando y los medios de contraste, si se usan, podrían entrar en el flujo sanguíneo fetal.
- Claustrofobia: es necesario que el paciente permanezca durante largos periodos de tiempo sin moverse dentro del equipo.

3.2.2. Ultrasonido (US)

El ultrasonido (US) no es un invento, sino un evento físico natural que puede ser provocado por el hombre. Siempre estuvo presente, sólo faltaban ojos observadores y mentes brillantes de personas en diferentes ramas de las ciencias para guiar su utilización, como ocurrió en el área de la medicina en donde produjo un gran impacto en el proceso diagnóstico. Su aplicación es el resultado de una serie de acontecimientos a lo largo de la historia, unidos a la perspicacia médica, curiosidad y habilidades de pioneros y sus continuadores en el campo de la investigación.

Actualmente, las aplicaciones del US mejoradas a través del tiempo se han desarrollado hasta llegar a ser una de las modalidades de diagnóstico más utilizadas, ya que es una herramienta de alta calidad, confiable y costo efectiva.

Su uso sin embargo, ha sido incorporado por más de una especialidad médica provocando controversia y a veces confusión. Nos preguntamos ¿Cómo mantener una adecuada utilización del US, bajo el manejo de médicos especialistas y/o tecnólogos acreditados dentro de un solo campo?

Una respuesta difícil si todo comenzó con una observación tres siglos atrás, y continuó con la participación de diferentes personajes y grupos, cuyos esfuerzos se han aunado para el bienestar de los pacientes y especialistas quienes finalmente cuentan con imágenes en tiempo real de la anatomía.

Desarrollo del US

A partir del siglo XVIII se hace notar el US como un fenómeno de la naturaleza cuando el biólogo italiano, Lazzaro Spallanzani descubre en el año 1700 la existencia de estas ondas, observando cómo los murciélagos atrapaban sus presas.

En la primera mitad del siglo XIX (1803-1853), el físico y matemático austriaco Christian Andreas Doppler presenta su trabajo sobre el .^Efecto Doppler.^observando ciertas propiedades de la luz en movimiento, que eran aplicables a las ondas del US. Sobre la base de este estudio los japoneses cien años más tarde desarrollarían lo que hoy conocemos como la aplicación del .^Efecto Doppler.^{en} US.

En la segunda mitad del siglo XIX los hermanos Pierre y Jacques Curie descubren las propiedades de algunos cristales conocidas como .^Efecto piezo-eléctrico", lo cual sirve de base para las diversas utilizaciones de las ondas de US.

A comienzos del siglo XX, se realiza una de las primeras aplicaciones en el área de la marina, después de que el físico francés Paul Langevin inventara el Sonar, en el cual se basó el posterior desarrollo de los equipos usados en la aviación y luego en medicina terapéutica y diagnóstica.

Terminada la segunda Guerra Mundial comienza el desarrollo de equipos diagnósticos en medicina, cuando grupos de investigadores japoneses, americanos y de algunos países europeos trabajan paralelamente para fabricar los primeros prototipos de equipos para diagnóstico médico en modo A (Analogue), y posteriormente en modo B (Bright) con imagen analógica (Figura 3.43).

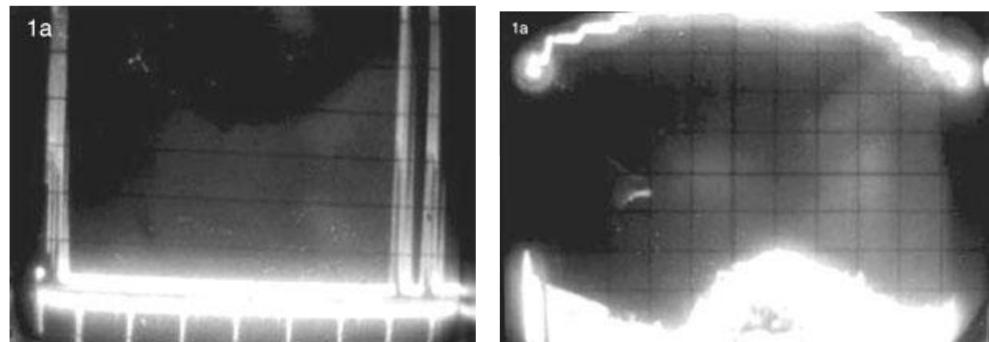


Figura 3.43: Primeras imágenes obtenidas con ultrasonidos de un quiste ovárico gigante.
a: Imagen en modo A. b: Imagen en modo B

Luego de varios años de desarrollo, en la década de 1950, el US es aceptado por las sociedades médicas como instrumento de diagnóstico en medicina, dando origen a un sinnúmero de trabajos de investigación en distintas áreas de aplicación. El primer artículo publicado en una revista científica de prestigio, fue en Lancet, en 1958, donde se describió la experiencia en un grupo de 100 pacientes normales y con patología abdominal (Figura 3.44).



Figura 3.44: Primer artículo publicado en una revista científica importante. Lancet, Junio 1958

En esta época los equipos eran de gran tamaño y ocupaban espacios considerables. No existía aún el gel conductor y los pacientes eran sumergidos en un estanque lleno con una solución conductora como el agua y debían permanecer sin moverse durante la adquisición de las imágenes (Figura 3.45).[23,24]

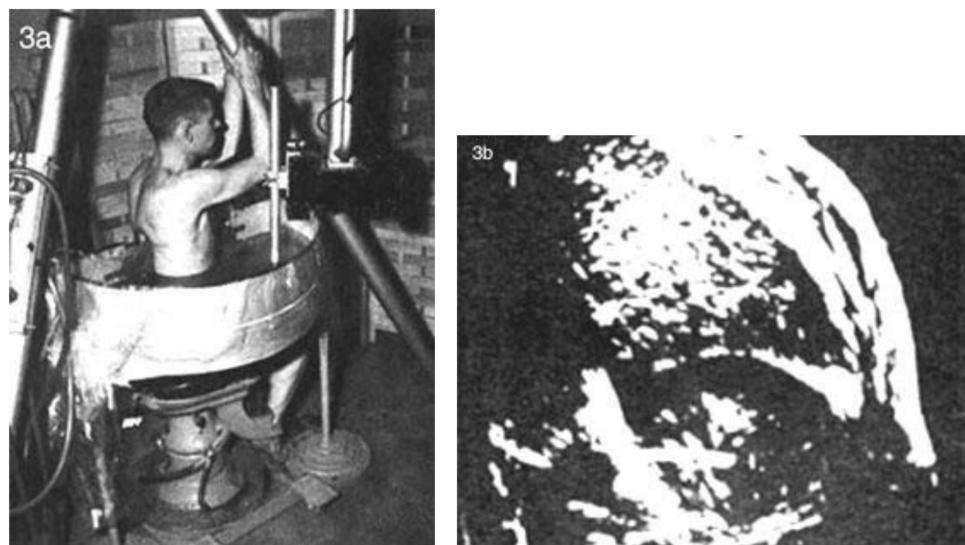


Figura 3.45: Año 1957. a: Uno de los primeros equipos utilizados en diagnóstico médico. b: Imagen del riñón derecho y parte del hígado obtenida con transductor que giraba alrededor del cuerpo sumergido en una batea con agua

A comienzos de la década de 1970 se introduce el "scan converter" el cual se logran las primeras imágenes de la anatomía en escala de grises. A fines de esta misma década, se agregan los micro-procesadores controlados, logrando finalmente imágenes en tiempo real de alta resolución. Así, progresivamente quedan obsoletas las máquinas estáticas con brazos articulados y lenta adquisición de las imágenes. La aceptación clínica es mayor y más rápida ya que existe una perspectiva real para el uso masivo de esta técnica inocua, de bajo costo y portátil.

Hacia fines de los años 70 y comienzos de los 80, además de los equipos estáticos, se utilizan los transductores mecánicos sectoriales.[25]

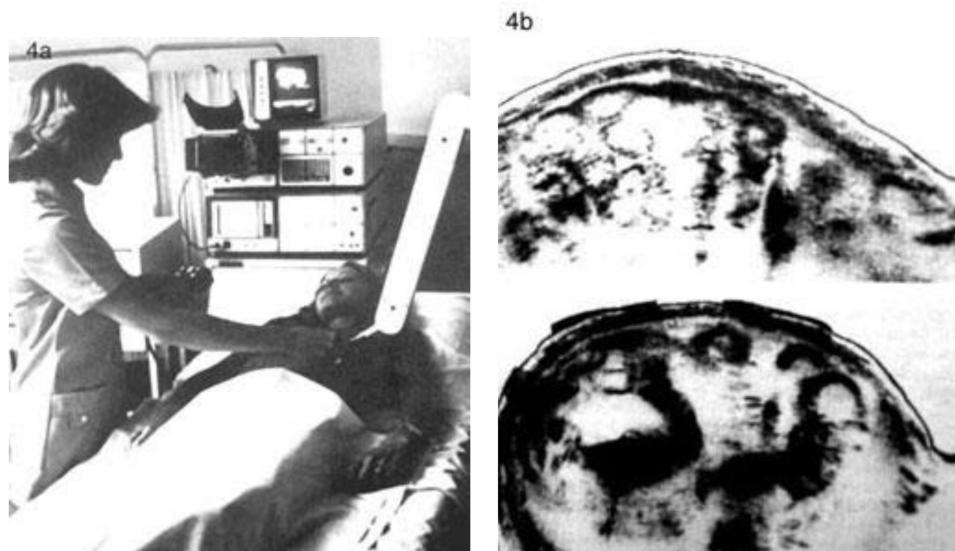


Figura 3.46: a: Equipo estático Picker del US. b: Imágenes del abdomen en Modo B

Física del Ultrasonido

Partiendo de lo básico, el Sonido se define como un tipo de energía mecánica que se propaga en ondas de presión sucesivas emitidas por una fuente con características elásticas y deformables que provocan la vibración de las moléculas de un objeto, por esto mismo podemos decir que el sonido es la percepción de que algo se está moviendo. Para que las ondas de energía puedan llegar a un objeto necesitan de un medio por el cual desplazarse, el cual se caracteriza por ser generalmente elástico. (Devesa, 2012)

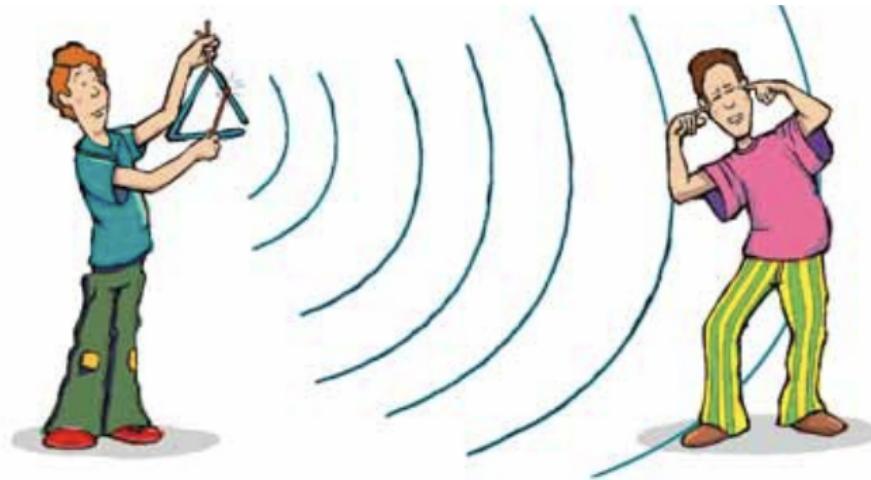


Figura 3.47: Ejemplo de propagación de onda

Como se mencionó anteriormente, el sonido se propaga en ondas, las cuales tienen características muy particulares, la primera que mencionaremos es el Ciclo, el cual es el fragmento de onda comprendido entre dos puntos iguales. (Díaz, 2012) La longitud de onda (λ) es la distancia entre dos fases de un ciclo y esta se puede medir en metros, centímetros o milímetros. (Díaz, 2012)

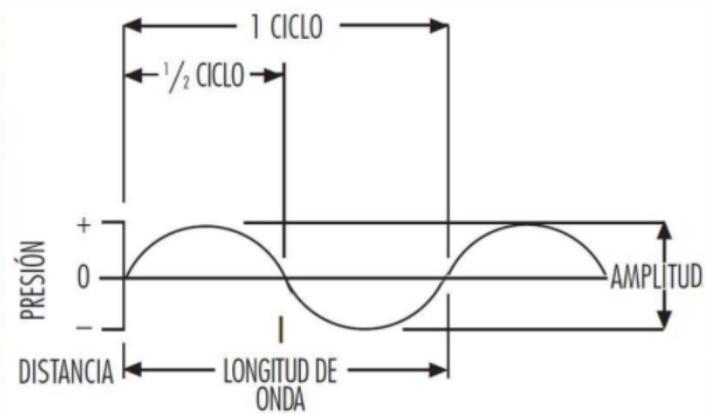


Figura 3.48

Otra de las características es la Frecuencia la cual se refiere al numero de ciclos o cambios

de presión que ocurren en un segundo. (Devesa, 2012)

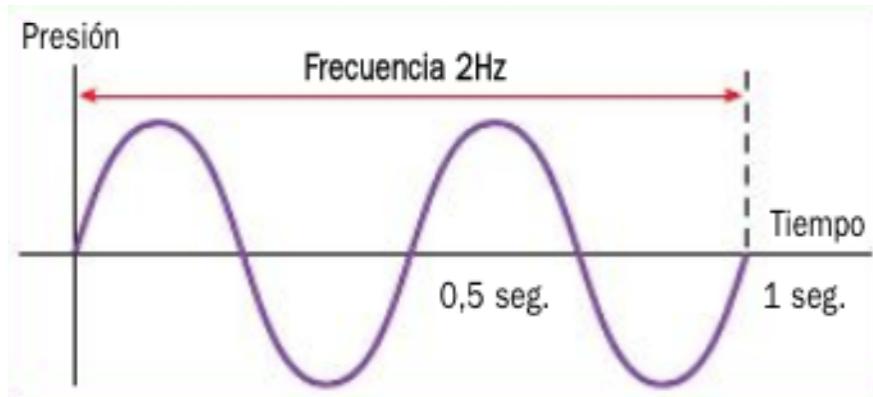


Figura 3.49

La unidad de medida de frecuencia es Hertzio (Hz), 1 Hertzio es equivalente a un ciclo cuyo recorrido dura 1 segundo.

$1 \text{ Hz} = 1 \text{ CICLO/SEGUNDO}$
$1 \text{ KhZ} = 1.000 \text{ CICLOS/SEGUNDO}$
$1 \text{ MHz} = 1.000.000 \text{ CICLOS/SEGUNDO}$
$1 \text{ GHz} = 1.000.000.000 \text{ CICLOS/SEGUNDO}$

Figura 3.50

Período es el tiempo en que tarda en producirse un solo ciclo, este tiene una relación inversa con la frecuencia. (Pineda , 2012)

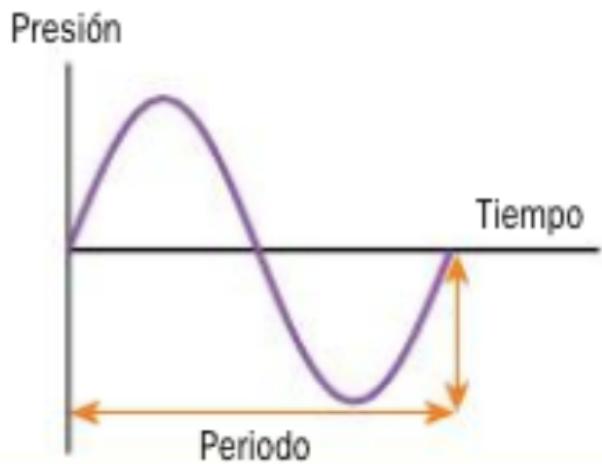


Figura 3.51

La **Amplitud de onda** se refiere a la altura máxima que alcanza desde su punto de origen, dependiendo de la altura nos determinará variables como la potencia, energía, electricidad, etc. (Devesa, 2012)

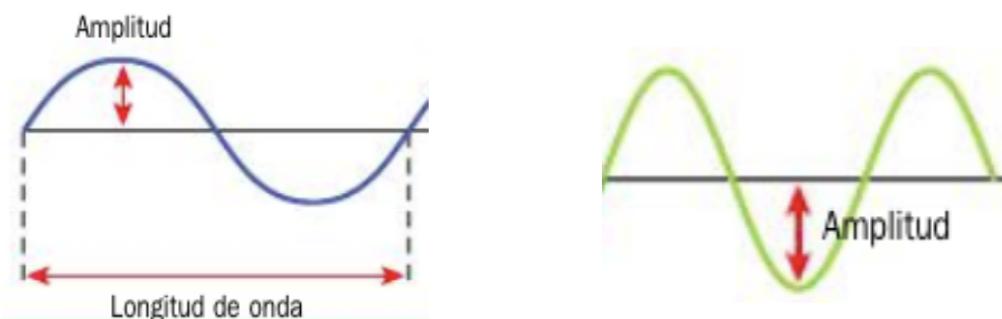


Figura 3.52

Por ultimo **Velocidad** es la distancia que recorre una onda sonora en la unidad de tiempo, la cual se calculará mediante la longitud de onda por la frecuencia, la cual se verá afectada al tener interacción con la materia ya que dependerá de la elasticidad (propiedad de regresar a su forma original), densidad (es igual al peso sobre el volumen del objeto), dureza (lo que carece de flexibilidad) y compresibilidad (capacidad de un cuerpo de modificar su volumen) principalmente. (Devesa, 2012)

Reflexión nos referimos a la cantidad de ondas de sonido que son emitidas y estas regresan a su punto de origen al que denominaremos eco, esta dependerá de la impedancia acústica de tejidos adyacentes así como longitud de onda del ultrasonido, características de superficie, extensión de la superficie y ángulo de incidencia del haz. (Devesa, 2012)

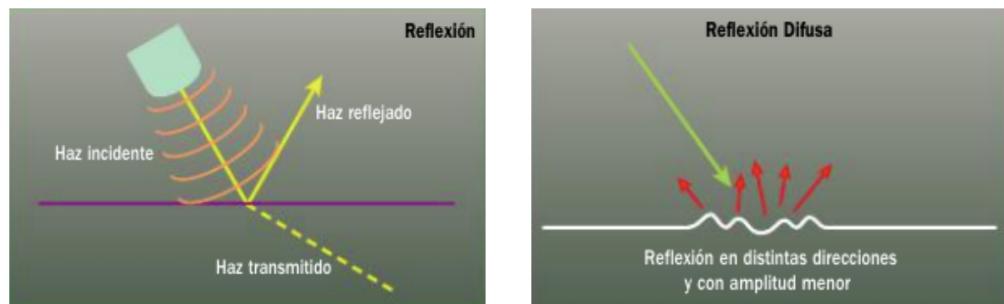


Figura 3.53

Refracción se refiere cuando el haz de ultrasonido sufre una desviación en su dirección al atravesar una superficie reflectante (como el agua) que separa 2 haz de ultrasonido con diferente velocidad de propagación. Si la superficie es muy irregular el haz se puede desviar tanto que no alcance ninguna estructura y lo veremos reflejado en nuestra pantalla como una imagen anicónica. (Devesa, 2012)

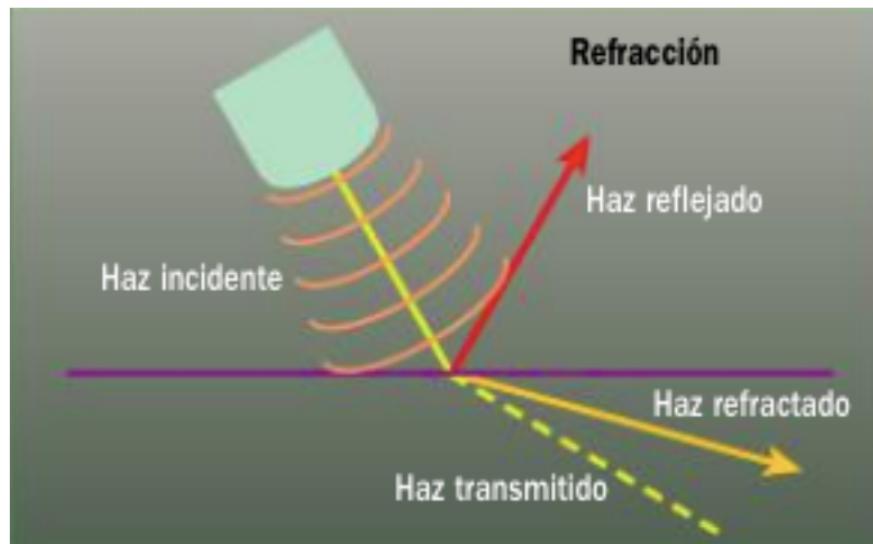


Figura 3.54

La **Absorción** se refiere al la energía del haz consumida al atravesar un medio Gsico, como ya mencionamos esa energía es traducida en calor al provocar la vibración de las partículas, este fenómeno se verá alterado por tres factores, frecuencia, viscosidad del medio y relajación de las moléculas. Por lo tanto las frecuencias altas provocan una absorción más alta teniendo una peor penetración del haz, por eso las frecuencias altas las utilizamos en el ultrasonido para valorar estructuras de tejidos blandos o superficiales, al contrario de las frecuencias bajas para valorar estructuras profundas abdominales ya que tienen una mejor penetración. (Devesa, 2012)

La **Atenuación** es la disminución de energía que sufre el haz de ultrasonido al atravesar el medio Gsico y esto estará directamente relacionado por los terminaños ya mencionados de reflexión, refracción, dispersión y absorción del medio. (Devesa, 2012) [26]

Teniendo en claro cada uno de estos conceptos básicos de Física del sonido sabemos que el ultrasonido utilizado como equipo biomédico genera ondas mecánicas originadas por la vibración de un cuerpo elástico como lo son los cristales **piezoelectrícios** los cuales se encuentran en nuestro traductor, que estos a su vez son activados por un impulso eléctrico provocando así transformación de la energía eléctrica a vibratoria y viceversa, las cuales se propagarán por medio de ondas a través de un medio físico como los tejidos, de tal manera que el transductor actúa como emisor y receptor de ultrasonidos para poder emitirnos una imagen en escala de grises en nuestro monitor.

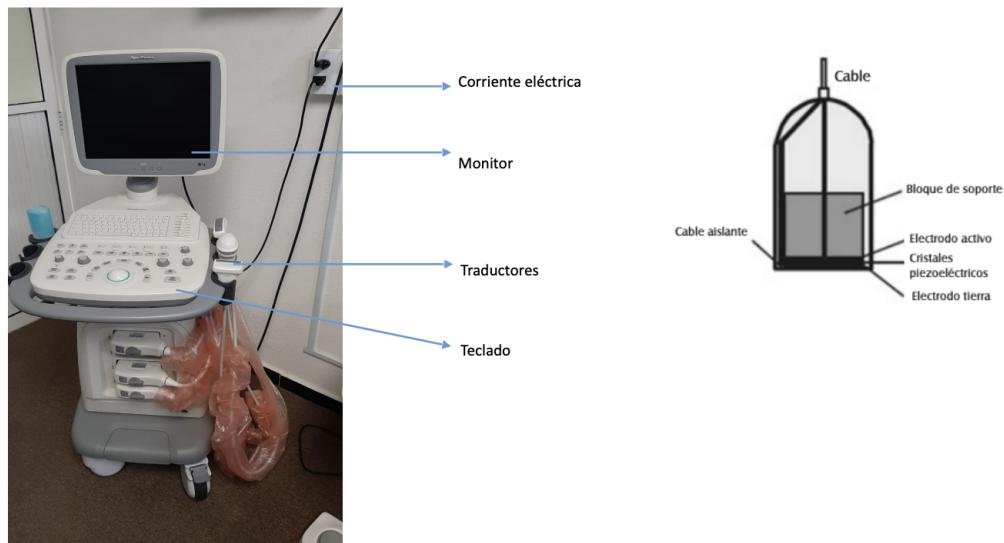


Figura 3.55



Figura 3.56

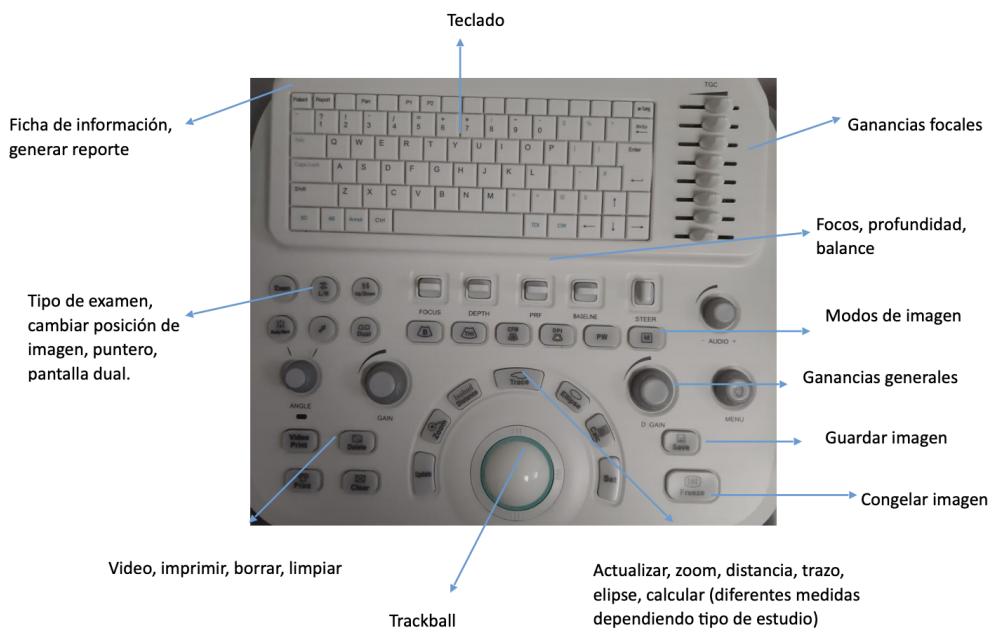


Figura 3.57

Los artefactos de imagen son ecos que aparecen en la imagen de nuestra pantalla pero que no corresponden en localización o intensidad con las estructuras exploradas del paciente.

Sombra acústica posterior: se forma detrás de una estructura que bloquea el paso de la onda de ultrasonido (calcio, aire, artefactos metálicos). (Rojas, 2017)



Refuerzo acústico posterior: detrás de una estructura anecoica que permite el paso de haz del ultrasonido (líquido). (Rojas, 2017)



Reverberación: ocurre al tener dos tejidos adyacentes con una diferencia notable en su impedancia acústica y altamente reflectantes donde se presenta un rebote repetitivo de la onda de ultrasonido entre las dos superficies, se producen imágenes similares repetitivas y equidistantes a la estructura real. (Rojas, 2017)



Figura 3.58

Cola de cometa: es un ejemplo de reverberación que se presenta cuando el haz de ultrasonido atraviesa interfaces próximas de muy diferente impedancia acústica, generando múltiples y pequeñas imágenes decrecientes, cortas; dando la apariencia de una cola de cometa. (Rojas, 2017)[27]

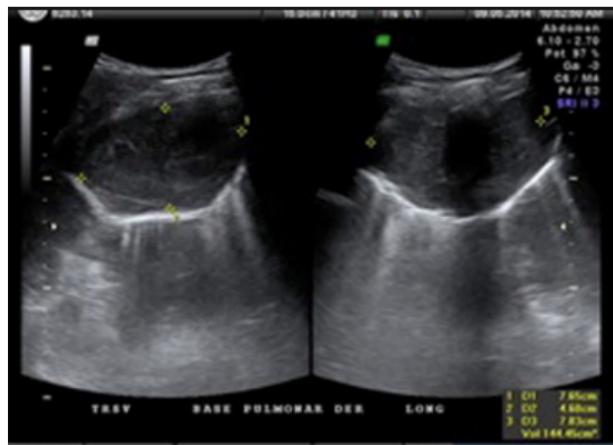


Figura 3.59

Imagen en espejo: es un tipo de reverberación que se forma cuando parte del haz del ultrasonido en su camino de retorno toma una ruta más larga o corta de regreso al transductor. El equipo emite el haz de ultrasonido en un tiempo de transmisión conocido, parte del haz atraviesa el tejido hasta atenuarse y cesar los ecos reflejados desde el mismo. La imagen en espejo se produce por la interacción del haz con una estructura altamente reflejante, que al entrar en contacto cambia su refracción por el tejido, generando ecos virtuales, simulando tratarse de estructuras más profundas, dando como resultado una imagen falsa en cuanto a la posición real de la estructura. No es frecuente sin embargo esto puede ocurrir al explorar la pleura, pelvis, tibia y el diafragma. (Rojas, 2017)[28]

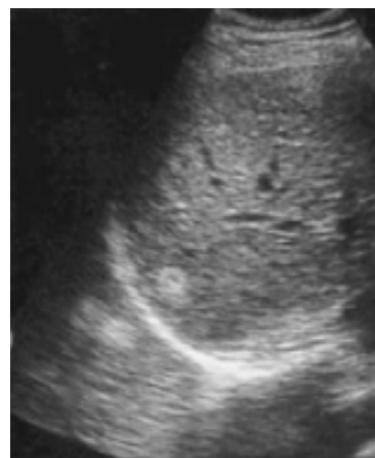


Figura 3.60

3.3. Estándares de Imagen Médica:

- **Tipos de Formato de Imagen:**

- DICOM
- NIFTI
- HL7

3.3.1. DICOM

Introducción

La atención sanitaria a la población se ha convertido en un objetivo prioritario en la sociedad actual. Las instituciones hospitalarias emplean a personal cada vez mejor formado, y también abundantes medios tecnológicos como sistemas PACS (Picture Archiving and Communication System) para permitir el correcto funcionamiento del estándar DICOM (**Digital Imaging and Communication in Medicine**) y poder así proporcionar un servicio más rápido y eficaz a los pacientes a través de plataformas informáticas. En este contexto, la interoperabilidad entre los diferentes sistemas de información y otros equipos utilizados en la atención de la salud de los ciudadanos es reconocida unánimemente como una de las claves para aumentar la calidad de la misma.

DICOM (Digital Imaging and Communications in Medicine) es un estándar propuesto y administrado por la National Electrical Manufacturers Association (NEMA). El propósito principal del estándar es garantizar la igualdad de condiciones desde el momento de la adquisición de un estudio imagenológico hasta el momento de ser desplegado en pantalla o impreso en papel radiográfico, después de un posible procesamiento de las imágenes.

Se debe resaltar la importancia que tienen las intensidades de grises puesto que estas deben ser las mismas sin importar el medio de visualización de la imagen, permitiendo que cualquier especialista observe lo mismo en cualquier tipo de dispositivo diagnóstico.

Aspectos Generales del Estándar DICOM

En la imagen 1 se aprecia una síntesis del concepto de DICOM siguiendo la estrategia de cartografía conceptual, cuyos contenido será analizado seguidamente (Figura 3.61).

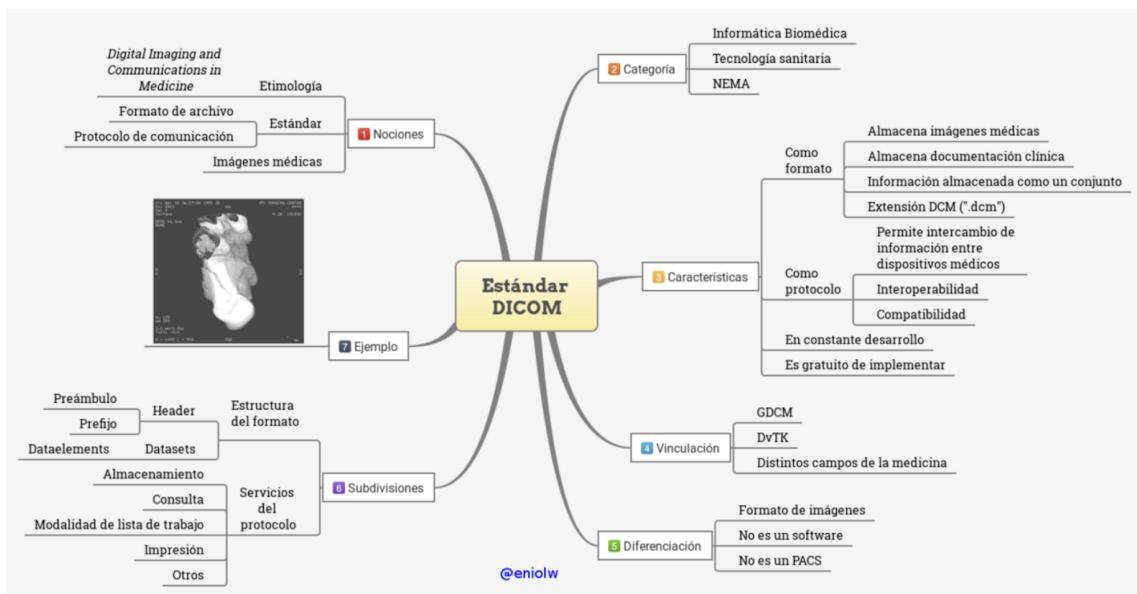


Figura 3.61: Cartografía conceptual sobre DICOM Autor: @eniolw Licencia: CC BY

DICOM es un acrónimo para “Digital Imaging and Communications in Medicine” que en español significa “**Imagen Digital y Comunicaciones en Medicina**”. Se trata de un estándar internacional que define un **formato de archivo** para representar información médica (imágenes y su metadata) y un **protocolo** para la comunicación y el correspondiente intercambio de información entre artefactos de uso clínico. Aunque se trata de una especificación que define ambos aspectos considerados, por lo general se le refiere simplemente como un formato.

Este estándar se puede **categorizar** como tecnología sanitaria dado que es estudiado y desarrollada para efectos de **informática biomédica**, que es la multidisciplina que aplica los conocimientos informáticos y médicos en la producción de tecnologías que contribuyan al mejoramiento de la salud.

Todo esto también pone de relieve que DICOM está vinculado con todas las áreas de la medicina donde tenga aplicación a través del uso de tecnologías para la generación de imágenes diagnósticas, llámese cardiología, oncología, ginecología, obstetricia, dermatología, radiología, neurología, oftalmología y muchas más.

Las características resaltantes de DICOM están en función de su modalidad. Como formato de archivo, DICOM permite la **representación y almacenamiento digital de imágenes médicas**. Estas imágenes abarcan varios tipos, como tomografías, ecografías, resonancias magnéticas, radiografías, angiografías, imágenes de medicina nuclear, entre

otras, pudiendo almacenar una o muchas imágenes en un único fichero. A continuación se muestra un esquema de los distintos tipos de imágenes médicas soportadas (Figura 3.62).[29]

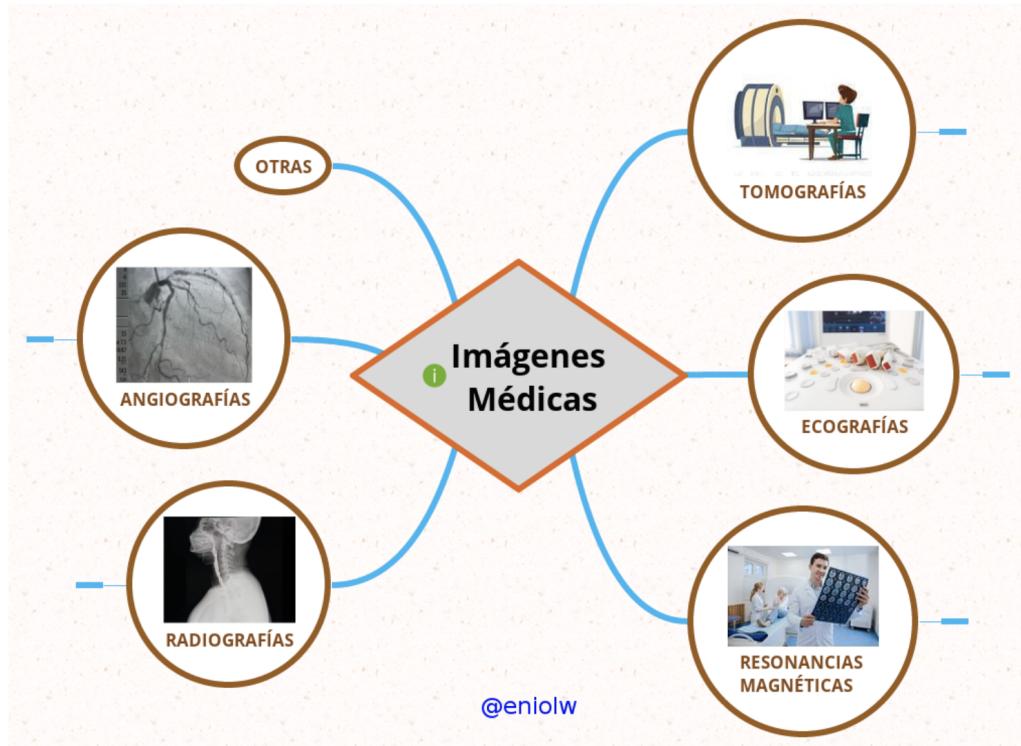


Figura 3.62: Imágenes médicas típicas Autor: @eniolw Licencia: CC BY

También, el formato almacena información de documentación clínica, lo que abarca una serie de datos sobre el paciente y el estudio, tales como nombre del paciente, su ID, fecha y hora del estudio, tipo de exploración, detalles sobre el equipo que tomó la imagen, distintas métricas y resultados, entre otros. La siguiente tabla muestra (apenas) algunos de estos datos categorizados (Tabla 1)

Tabla 1: Algunas etiquetas de un archivo DICOM	
Paciente	
Nombre, ID,	
Fecha de nacimiento, Hora de nacimiento	
Lengua materna	
Dirección de habitación, Número telefónico	
Grupo étnico	
Ocupación	
Visita	
Nombre, Dirección, Código	
Nombre del médico que refiere	
Responsable de la organización	
Operador	
Estudio	
ID	
Fecha programada de inicio	
Médico solicitante	
Motivo del estudio	
Fecha culminación	
Comentarios	
Interpretación	
Autor	
Descripción del diagnóstico	
ID del estatus	
Equipo	
Fabricante	
Nombre de la estación	
Versión del software	

Algo resaltante y característico del formato DICOM es el hecho de que esta información es almacenada conjuntamente, es decir, la documentación clínica es parte de las imágenes médicas y viceversa. La idea es que las imágenes sean indivisibles de su respectiva información, sean portadas juntas y no sean separadas por error, de tal modo que los programas habilitados para leer DICOM correctamente, podrán abrir toda esta data y básicamente mostrarla al personal encargado de interpretarla.

Aunque otros formatos siguen un patrón similar a DICOM en el sentido de que soportan el almacenamiento de imágenes y metadatos (como JPG), **DICOM se diferencia de ellos** puesto que provee un registro de datos más extenso y sofisticado, con múltiples atributos/campos para ser llenados por la aplicación que escriba DICOM. Además, DICOM puede no solamente almacenar una imagen, sino una colección de ellas a modo de “*frames*”. El formato también permite que las imágenes almacenadas puedan tener distintos tipos de compresión, incluyendo las de JPG.

A su vez, como protocolo, DICOM se caracteriza por permitir el **intercambio de información entre distintos dispositivos y sistemas de propósito biomédico**, tales como las decenas de equipos para la toma de imágenes médicas, estaciones de trabajo diagnóstico, servidores de red, accesorios, PACS (sistema de archivado y comunicación de imágenes), Sistemas de Información de Radiología (RIS), Sistemas de Información de Cardiología (CIS), impresoras, etc.

Esta integración es, de hecho, una de las motivaciones principales por las que surgió DICOM. Ante la aparición de muchos equipos biomédicos heterogéneos y provenientes de distintos fabricantes, también aparecían obstáculos ocasionados por su incompatibilidad e incapacidad de estos para trabajar juntos. De allí que el surgimiento del estándar viene a suplir de **integración y compatibilidad** para poder transmitir, almacenar, consultar, imprimir, procesar y visualizar la información de las imágenes médicas entre todos estos equipos y sistemas. La imagen 3 es una representación del trabajo conjunto de distintos hardwares y sistemas biomédicos trabajando conjuntamente gracias a DICOM (Figura 3.63).[30]

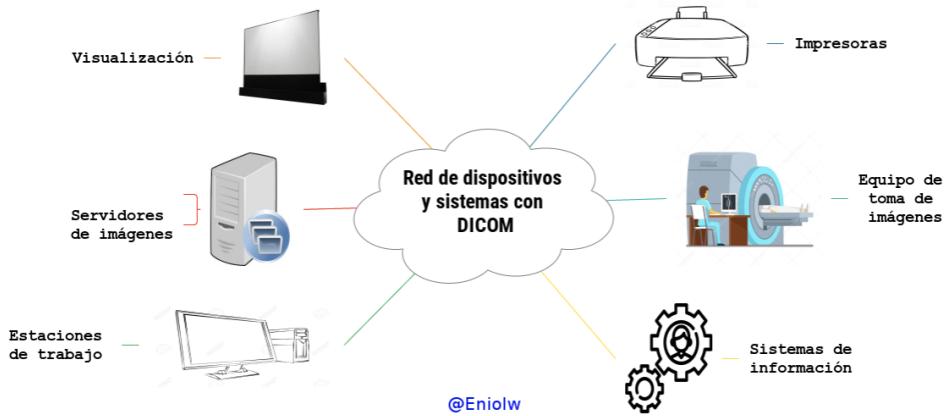


Figura 3.63: Sistemas y dispositivos de uso médico Autor: @eniolw CC BY

Ahora bien, es importante destacar que DICOM no es un software. No es un programa que se descarga, instala y usa, sino que es un estándar, es decir, una especificación, un plano que define las reglas como debe ser almacenada digitalmente una data y cómo deben comportarse las aplicaciones de distintos dispositivos a la hora de trabajar conjuntamente con esa data. Por tanto, el software sería aquel que implementando el estándar en su funcionamiento, abre el archivo en formato DICOM, lo visualiza y/o aplica cualquiera de los varios servicios soportados.

De hecho, es posible que la mayoría de los usuarios no dispongan en su computadora un software que pueda abrir apropiadamente DICOM y puede que el sistema operativo ni siquiera reconozca la extensión ".dcm", amén de que el formato es algo más especializado. Algunos lectores o visores DICOM conocidos son Dicompyler, Open Dicom Viewer, Microdicom, OsiriX, RadiAnt, entre otros, aunque también hay servicios online que abren archivos DICOM y los visualizan a través del navegador web. La imagen 4 contiene un ejemplo de visor DICOM (Figura 3.64). En ocasiones, los radiólogos almacenan los archivos DICOM en CD-ROMs para archivado o traslado, y se incluye allí al software necesario para desplegarlos.

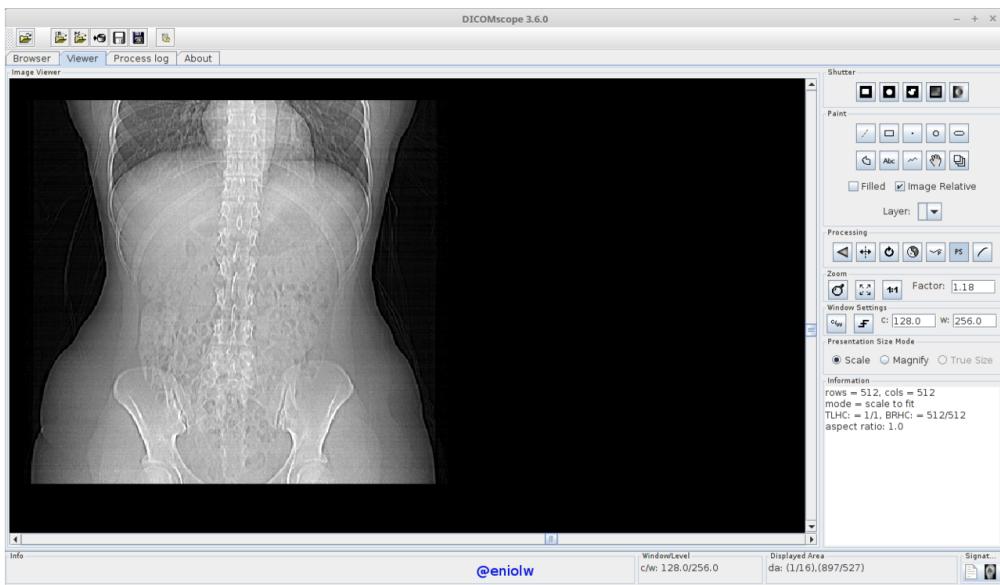


Figura 3.64

DICOM está en **constante desarrollo** para mantenerse actualizado con el surgimiento de más novedosos equipos y sistemas, que se actualizan tan rápido como la medicina misma. De hecho, un vistazo a la página oficial del estándar revela una cronología que data de 1980 registrando no menos de 21 versiones mayores del formato, siendo la última la del año 2017. También, la organización que lo regula ha dejado en claro que DICOM, en tanto estándar, está pensado para ser **implementado gratuitamente** por cualquier fabricante.[31]



Figura 3.65: Imagen médica en DICOM Fuente: DicomLibrary.com

Aspectos Técnicos de DICOM

Desde un punto de vista más técnico, DICOM en todas sus versiones está definido a lo largo de más de 20 documentos oficiales, que comprenden Vista General, Conformación, Definición de Objetos de Información, Especificaciones de Clases de Servicios, Estructura de Datos y Codificación, Diccionario de Datos, Intercambio de Mensajes, Soporte de Comunicación en Red para Intercambio de Mensajes, Almacenamiento de Medios y Formatos de Archivo para el Intercambio de Mensajes, Perfiles de Aplicaciones de Almacenamiento de Medios, Formatos de Medios y Medios Físicos para el Intercambio de Imágenes, Función de Visualización Estándar de Escala de Grises, Perfiles de Seguridad y Administración de Sistemas, Recursos de Mapeo de Contenido, Información Exploratoria, Servicios Web, Hosting de Aplicaciones, Reporte de Imágenes usando la Arquitectura de Documentos Clínicos HL7 y Transformaciones entre DICOM y otras Representaciones.

Ahondando un poco en lo que refiere a la Estructura de Datos y Codificación y la Definición de Objetos de Información, se tiene que en cuanto a las *subdivisiones*, la estructura del formato DICOM es algo elaborada. A grandes rasgos consta de un “encabezado” (*header*) y un “conjunto de información” (*dataset*) [32].

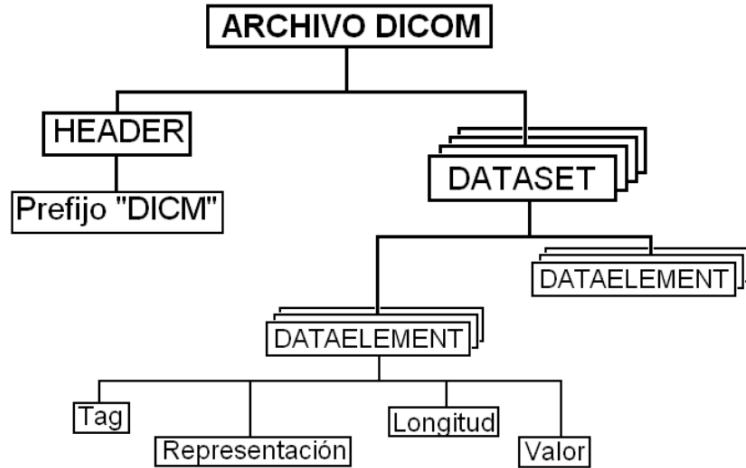


Figura 3.66: Estructura de un Archivo DICOM

El encabezado puede contener algunos atributos y especificaciones sobre la aplicación que escribe el fichero. Este encabezado posee dos secciones: un preámbulo (*preamble*) de 128 bytes que almacena información sobre una miniatura del fichero, además de proporcionar compatibilidad con otros formatos, y un prefijo (*prefix*) de 4 *bytes* que únicamente registra los caracteres “DICM”.

Por su parte, el conjunto de información (*dataset*) está destinado a incluir toda información médica pertinente. El dataset está organizado siguiendo el Modelo de Información de DICOM, esto es, una serie de clases (concepto de programación orientada a objetos) que modelan toda la información, destacándose el “Par Objeto-Servicio” (*Service-Object Pair SOP*). Este par consta de “Elementos de servicio DICOM” (*DICOM Service Elements DIMSE*) vinculados con una “Definición de Objetos de Información” (*Information Object Definition IOD*). Es decir, según esto, DICOM consiste en una serie de servicios aplicados sobre atributos.

Los IOD consisten en una serie de atributos que son codificados a manera de “Elementos de información” (*dataelements*). Cada *dataelement* se conforma por una etiqueta de identificación (*Tag*), una representación de valor (*Value Representation VR*) que señala el tipo de dato (entero, flotante, carácter, etc.) del contenido a registrar; la longitud del valor (*Length Representation LR*) y, por supuesto, el campo de valor (*Value Field VF*), que almacena el contenido de la información útil (nombre, edad, estudio realizado, la imagen médica, etc.). Cada etiqueta es un identificador único para cada Data Element compuesto de dos partes. El DICOM utiliza la notación siguiente para referir a una etiqueta: (gggg,eeee) el primer valor hexadecimal de 2 bytes es el número de grupo y el segundo es el número del elemento. Por ejemplo (0010,0030) corresponde a la fecha de nacimiento del

paciente. La imagen 6 es una representación gráfica estandarizada de esto (Figura 3.67). Se podrá conseguir más información sobre la estructura y modelo de datos de DICOM en las referencias.

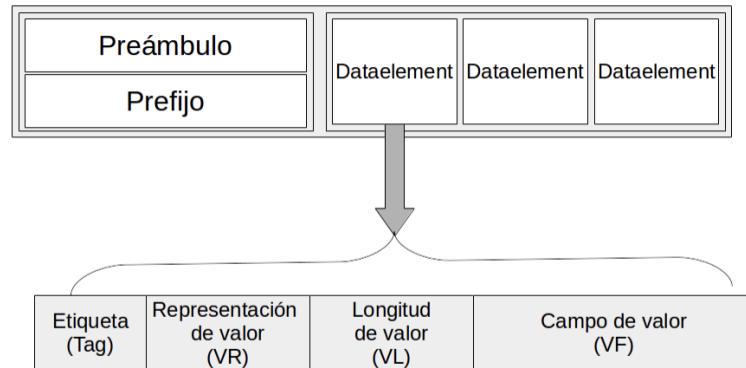


Figura 3.67: Estructura interna de un archivo DICOM Autor: @eniolw CC BY

Existe una gran variedad de Data Elements y no siempre estarán definidos en su totalidad dentro de un archivo, asimismo habrán Data Elements que no aporten información relevante para ciertas necesidades. Por esto es importante saber que como mínimo el archivo debe contener los Data Elements mencionados en la Figura 3.66 para la adecuada lectura de una imagen:

TAG	Descripción	Tipo
(0028,0002)	Samples per Pixel	Int
(0028,0008)	Number of Frames	Int
(0028,0010)	Rows	Int
(0028,0011)	Columns	Int
(0028,0100)	Bits Allocated	Int
(0028,0101)	Bits Stored	Int
(0028,0102)	Hight Bit	Int
(0028,0103)	Pixel Representation	Int
(0028,1050)	Window Center	Int
(0028,1051)	Window Width	Int
(0028,1052)	Rescale Intercept	Int
(0028,1053)	Rescale Slope	Int
(7FE0,0010)	PixelData	Byte[] ó Uint16[]

Figura 3.68: Tabla de Data Elements Necesarios para la lectura adecuada de una imagen

La manera como están escritos los Data Element se denominan Sintaxis de Transferencia, que generalmente es igual para todos los elementos de un archivo. La sintaxis de transferencia determina si el Data Element está escrito en un ordenamiento Big o Little Endian, si el valor de representación está o no incluido en el elemento (VR explícito o Implícito) y el tipo de compresión de la imagen (Mapa de Bits, JPEG o diferentes tipos de compresión).

Los datos del píxel se pueden enviar en un formato nativo o en un formato encapsulado definido fuera del estándar DICOM. Si están enviados en un formato nativo, las muestras del píxel se codifican como el encadenamiento directo de los bits de cada muestra del píxel. Si los bits están almacenados en un formato encapsulado, las muestras del píxel se codifican según el proceso de codificación definido por uno de las sintaxis de transferencia.

Los servicios DICOM abarcan Store, que envía la información persistente a un PACS o estación de trabajo para ser “almacenada”; *Retrieve*, mediante el cual un equipo puede “consultar” en un PACS listados de objetos DICOM y seleccionarlos; *Modality worklist*, que gestiona una “lista de procedimientos” para la generación de imágenes a ser ejecutada por un PACS o similar; *Print*, con el cual se envían los objetos DICOM como órdenes de “ impresión” a las impresoras, entre otros.

Por otra parte, DICOM tiene **vinculación** con GDCM, siglas de *Grassroots DICOM* (bases de DICOM), que es la más difundida implementación del estándar. Es una librería multiplataforma escrita en lenguaje C++, en software libre y código abierto ideada para trabajar con archivos DICOM. Ofrece soporte para otros lenguajes como Python, Java, C#, PHP y Perl a través de respectivos *wrappers* (librerías intermedias o “envoltorios”).

GDCM es la base de varios proyectos de software como Gdcmviewer, Gdcmanon (para anonimizar la data), Gdemimg (para manipular los elementos de datos de pixel de los archivos DICOM), entre muchos otros, algunos de las cuales se listan en la wiki oficial de la librería.

Con respecto al lenguaje Python, la librería ImageIO también tiene capacidad para el procesamiento de gráficos volumétricos y de interés científico, entre lo cual se cuenta DICOM. Sin embargo, una librería en Python más sofisticada para esto es Pydicom. Como se lee en su documentación oficial:

“Pydicom es un paquete de Python puro para trabajar con archivos DICOM tales como imágenes médicas, reportes y objetos de radioterapia. Pydicom facilita el leer estos complejos archivos en estructuras pythónicas naturales para una sencilla manipulación. Los conjuntos de datos modificados pueden ser escritos de nuevo en archivos de formato DICOM”

Esta librería tiene potencial para muchos tipos de aplicaciones. A grandes rasgos, con ella se puede:

1. Leer el núcleo de elementos de DICOM.
2. Escribir nuevos archivos DICOM.
3. Trabajar con Pixel Data.
4. Manejar data de imágenes comprimidas.
5. visualizar imágenes DICOM.

3.3.2. NIFTI

El formato de archivo de la “**Neuroimaging Informatics Technology Initiative**” (NIFTI) se imaginó hace aproximadamente casi dos décadas como un reemplazo del formato de archivo 7.5, entonces extendido pero problemático. El principal problema con el formato anterior fue quizás la falta de información adecuada sobre la orientación en el espacio, de modo que los datos almacenados no pudieran ser interpretados sin ambigüedades. Aunque el archivo fue utilizado por muchos softwares de imágenes diferentes, la falta de información adecuada sobre la orientación obligó a algunos a incluir para cada archivo de análisis, un archivo adjunto que describa la orientación, como un archivo con extensión .mat. El nuevo formato se definió en dos reuniones del denominado Grupo de Trabajo de Formato de Datos (dfwg) en los Institutos Nacionales de Salud (nih), uno el 31 de marzo y otro el 02 de septiembre de 2003. Representantes de algunos de los programas de neuroimagen más populares acordaron un formato que incluiría nueva información, y al usar el nuevo formato, ya sea de forma nativa, o tenerlo como opción para importar y exportar.

Quizás la consecuencia más visible de la falta de información de orientación fue la confusión reinante entre los lados izquierdo y derecho de las imágenes cerebrales durante los años en que el formato de análisis era dominante. Fue en este momento cuando los investigadores se acostumbraron a describir una imagen como parte de una convención "neurológica.^o radiológica". Estos términos siempre han sido inadecuados porque, en ausencia de información de orientación, no necesariamente dos piezas de software tendrían que mostrar el mismo archivo con el mismo lado del cerebro en el mismo lado de la pantalla. Un archivo puede mostrarse en orientación "neurológica.^{en} un software, pero en orientación radiológica en otro. El formato nifti evitó todos estos problemas, volviendo estos términos obsoletos.

El software ahora puede marcar el lado izquierdo o derecho correctamente, a veces dando la opción de mostrarlo invertido para adaptarse mejor a la preferencia de orientación personal del usuario.

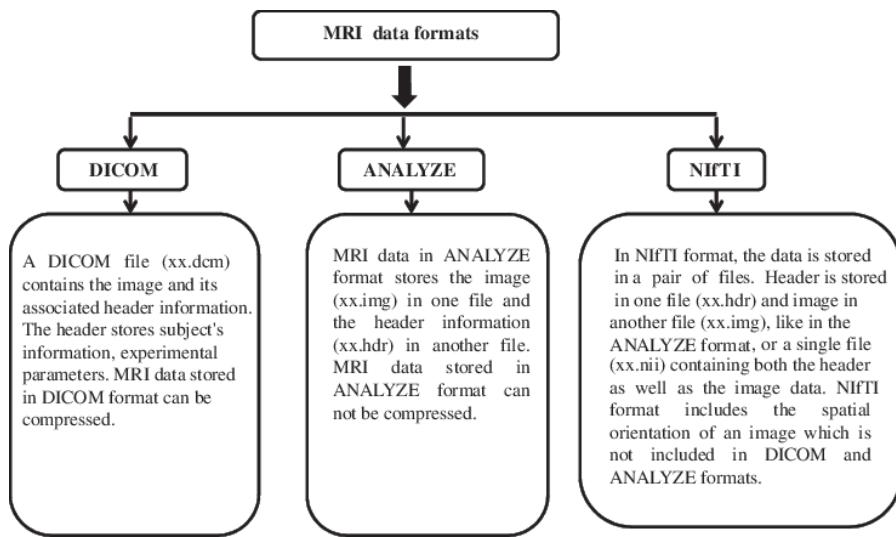


Figura 3.69

Mismo Formato, Diferentes Presentaciones

Una sola imagen almacenada en el formato de ANALYZE 7.5 requiere dos archivos: un **encabezado**, con extensión **.hdr**, para almacenar metainformación, y los datos reales, con extensión **.img**. Para mantener la compatibilidad con el formato anterior, los datos almacenados en formato nifti también usan un par de archivos, **.hdr / .img**. Se tuvo cuidado para que la estructura interna del formato nifti fuera mayormente compatible con la estructura del formato de ANALYZE. Sin embargo, el nuevo formato agregó algunas mejoras inteligentes. Trabajar con un par de archivos para cada imagen como en **.hdr / .img**, en lugar de solo uno, no solo es inconveniente, sino que también es propenso a errores, ya que uno podría olvidar fácilmente (o no saber) que los datos de interés en realidad se divide en más de un archivo.

Para abordar este problema, el formato nifti también permite el almacenamiento como un solo archivo, con extensión **.nii**. Sin embargo, un solo archivo o un par de archivos no son las únicas presentaciones posibles. Es muy común que las imágenes tengan grandes áreas de fondo sólido o archivos que describan máscaras y regiones de interés que contienen solo unos pocos valores únicos que aparecen repetidos muchas veces. Los archivos como estos ocupan un gran espacio en el disco, pero con poco contenido de información real. Este es el caso perfecto donde la compresión puede lograr excelentes resultados. De hecho, tanto archivos nifti como ANALYZE se pueden comprimir. El algoritmo de DEFLATE (utilizado, por ejemplo, por gzip) puede funcionar en flujos, permitiendo la compresión y descompresión sobre la marcha. Las versiones comprimidas tienen la extensión **.gz** agregada: **.nii.gz** (archivo único) o **.hdr /.img.gz** (par de archivos, ya sea nifti o analizar).

THE NIFTI1 DATA FORMAT



Figura 3.70

Dimensiones Predefinidas para Espacio y Tiempo

En el formato nifti, las tres primeras dimensiones están reservadas para definir las tres dimensiones espaciales - **x**, **y** y **z** -, mientras que la cuarta dimensión está reservada para definir los puntos de tiempo - **t**. Las dimensiones restantes, del quinto al séptimo, son para otros usos. Sin embargo, la quinta dimensión todavía puede tener algunos usos predefinidos, como almacenar parámetros de distribución específicos de voxel o mantener datos basados en vectores.

Un archivo NIfTI para una exploración MRI típica almacena los valores de voxel en una matriz de números. Las coordenadas para un solo voxel dentro de un volumen de imagen NIfTI se pueden especificar como un índice tridimensional (x, y, z) o un índice cuatro dimensiones (x, y, z, t). En el ejemplo a continuación, el voxel ubicado en el archivo nifti en el índice (x, y, z) = (127, 259, 368) tiene el valor de intensidad 2250 (unidades arbitrarias).

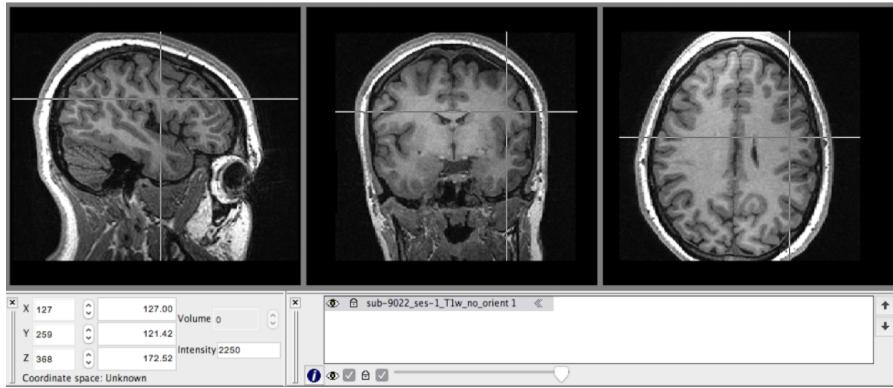


Figura 3.71

Información de Adquisición del Corte

The fields **char slice_code**, **short slice_start**, **short slice_end** and **float slice_duration** are useful to store information about the timing of an fMRI acquisition, and need to be used together with the **char dim_info**, which contains the field **slice_dim**. If, and only if, the **slice_dim** is different than zero, slice_code is interpreted as:

CODE INTERPRETATION	
0	Slice order unknown
1	Sequential, increasing
2	Sequential, decreasing
3	Interleaved, increasing, starting at the 1st MRI slice
4	Interleaved, decreasing, starting at the last MRI slice
5	Interleaved, increasing, starting at the 2nd MRI slice
6	Interleaved, decreasing, starting at one before the last MRI slice

Figura 3.72

Los campos **short slice_start** y **short slice_end** informan, respectivamente, cuáles son los primeros y los últimos sectores que corresponden a la adquisición de resonancia magnética real. Los sectores presentes en la imagen que están fuera de este rango se tratan como sec-

tores rellenos (por ejemplo, que contienen ceros). El campo float slice_duration indica la cantidad de tiempo necesaria para adquirir un solo segmento. Tener esta información en un campo separado permite almacenar correctamente imágenes de experimentos en los que slice_duration * dim [slice_dim] es menor que el valor almacenado en pixdim, generalmente el tiempo de repetición (tr).

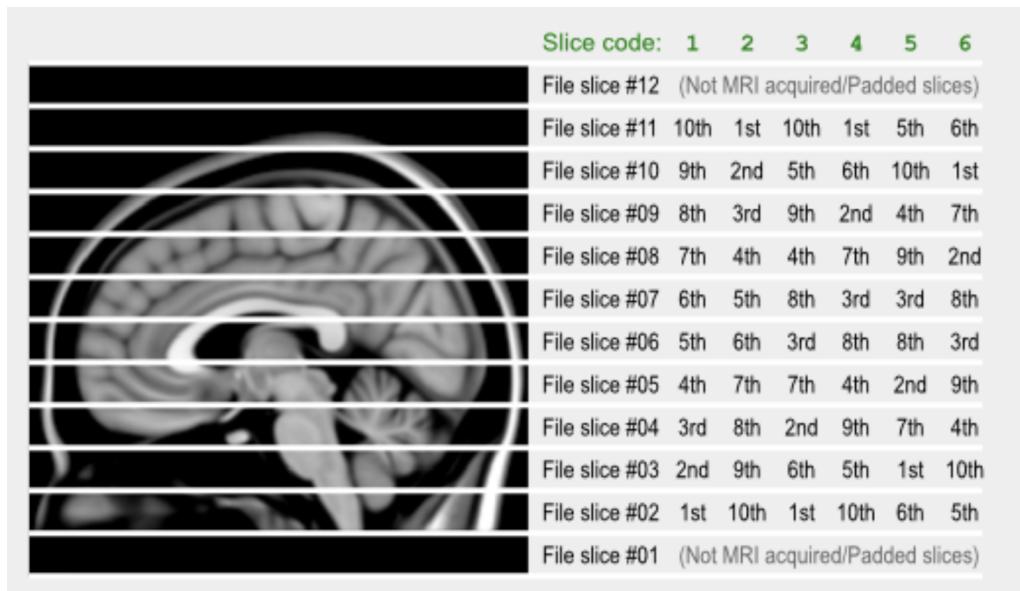


Figura 3.73

Problemas

El formato nifti trajo una serie de grandes beneficios si se compara con el antiguo formato de análisis. Sin embargo, también trajo su propio conjunto de nuevos problemas. Afortunadamente, estos problemas no son graves. Aquí están algunos:

- Aunque se hizo un gran esfuerzo para mantener la compatibilidad con el análisis, no se preservó un aspecto crucial: se supone que el sistema de coordenadas del mundo, en el formato nifti, es ras, lo cual es extraño y confuso. La última es una opción mucho más lógica desde una perspectiva médica. Afortunadamente, dado que la orientación se almacena sin ambigüedades, es posible voltear las imágenes en la pantalla a voluntad en la mayoría de los programas.
- El formato de archivo todavía depende demasiado de que la extensión del archivo sea .nii o de un par .hdr / .img, en lugar de cadenas o números mágicos mucho menos ambiguos. Por otro lado, las diferentes cadenas mágicas para un solo archivo y para pares de archivos evitan efectivamente la posibilidad de división / fusión de archivos

utilizando herramientas comunes de sistema operativo (como dd en Linux), ya que la cadena mágica debe cambiarse, aunque la estructura del encabezado sigue siendo absolutamente idéntica.

- La cadena mágica que está presente en el encabezado no se coloca al principio, sino cerca de su final, lo que hace que el archivo sea prácticamente irreconocible fuera del campo de neuroimagen.
- No todos los paquetes de software implementan el formato exactamente de la misma manera. Los datos basados en vectores, por ejemplo, que deben almacenarse en la quinta dimensión, a menudo se almacenan en la cuarta, que debe reservarse por tiempo. Aunque esto no es un problema con el formato en sí, pero con el uso que se hace de él, las prácticas incorrectas de implementación fácil conducen a la difusión de archivos ambiguos y mal formados que eventualmente no se pueden leer en otras aplicaciones según lo previsto en el momento de la creación del archivo.

A pesar de estos problemas, el formato ha tenido mucho éxito como un medio para intercambiar datos entre diferentes paquetes de software. La versión actualizada nifti 2.0 retiene generalmente la misma cantidad de información que la anterior, con la diferencia crucial de que permite muchos más puntos de datos en cada dimensión, lo que permite que se use la misma estructura general de archivos para almacenar, por ejemplo, datos escalares basados en la superficie, o grandes matrices de conectividad.

Ninguno de estos se había previsto originalmente en el momento en que se desarrollaron los formatos de análisis o nifti-1. [33]

3.3.3. HL7

Health Level Seven (HL7) - *Nivel Siete en Salud*, por sus siglas en inglés- proporciona estándares de interoperabilidad que mejoran el intercambio electrónico de información médica, optimizan el flujo de trabajo y reducen la ambigüedad. Su sintaxis coherente y extensible permite estructurar información en salud, apoyar los procesos de atención al paciente, para ser intercambiada entre aplicaciones de software, conservando al mismo tiempo la semántica de la información en todas las plataformas. Además brinda flexibilidad, porque es posible desarrollar aplicaciones en diferentes entornos tecnológicos y conectarlas entre sí.[34,35]

HL7 History

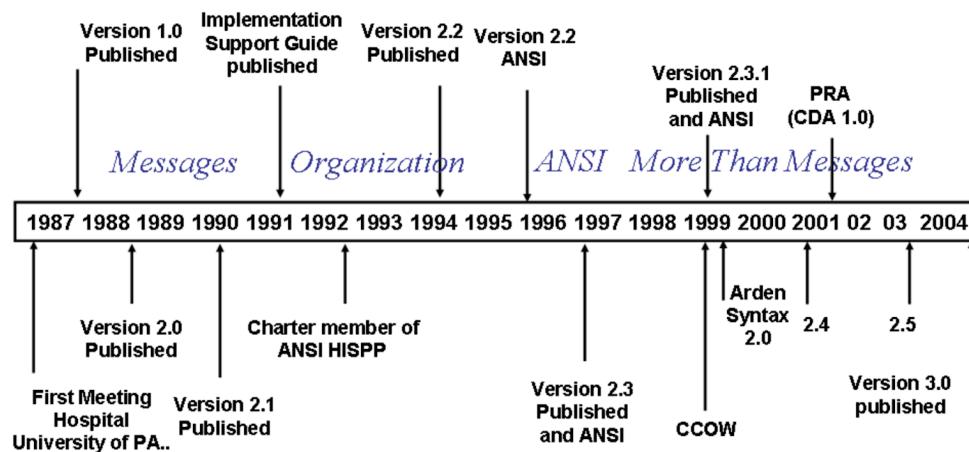


Figura 3.74

Estos estándares son acreditados por la organización **American National Standards Institute (ANSI)** - *Instituto Nacional Estadounidense de Estándares*, por sus siglas en inglés-. Desde su origen, el nombre de HL7 se asociaba a las versiones del estándar de mensajería para el intercambio electrónico de datos de salud, el cual está enfocado al intercambio de datos entre aplicaciones desarrolladas por diferentes fabricantes, facilitando el desarrollo de interfaces para la comunicación entre los sistemas de información médica, como son los RIS y los PACS. Sin embargo, la creciente necesidad de generar sistemas de información integrados regionalmente, hizo necesario el desarrollo de un espectro más amplio de estándares que faciliten la interoperabilidad.⁷ Por esta razón, a partir del año 2000 la organización HL7 cuenta con un proceso para definir una serie de herramientas de interoperabilidad, esto dio origen a que a partir del 17 de diciembre de 2003, la ANSI aprobara una serie de estándares contenidos en la especificación de la versión 3 HL7.

Esta versión no solo aporta una nueva arquitectura para establecer los modelos de mensajes, sino que incorpora una completa metodología para definir, desarrollar e implantar esquemas de interoperabilidad basados en mensajes y documentos. HL7 v3 está construida a partir de conocidos estándares de la industria, como son Unified Modeling Language (UML) -Lenguaje Unificado de Modelado, por sus siglas en inglés- y Extensible Markup Language (XML) -Lenguaje de Marcas Extensible, por sus siglas en inglés-. Todos sus elementos estructurales proceden de un núcleo compacto denominado Reference Information Model (RIM) -Modelo de Referencia de Información, por sus siglas en inglés-, lo que

facilita su adaptabilidad y reusabilidad.⁷ A partir de este marco, HL7 v3 define escenarios reales de interoperabilidad para cada dominio asistencial (admisión, laboratorio, radiología), que incluyen los roles de las aplicaciones participantes (emisor - receptor), eventos activadores, dinámica de la interacción de objetos y la configuración de mensajes. Uno de los estándares contenidos dentro de esta nueva versión está relacionado con la arquitectura de documentos clínicos electrónicos, este recibe el nombre de **Clinical Document Architecture (CDA)** - *Arquitectura de Documentos Clínicos*, por sus siglas en inglés.

Servicios HL7 CDA

HL7 CDA, es un estándar basado en XML para el marcaje de documentos, que puntuiza la estructura y semántica de documentos clínicos con el propósito de facilitar su intercambio en un entorno de interoperabilidad. Este estándar utiliza los elementos de HL7 v3 (RIM, Tipos de Datos, XML), para definir la estructura y la semántica de documentos clínicos. Como resultado de esta plataforma común, se puede implementar esquemas de interoperabilidad de datos sanitarios basados en mensajería HL7 v3 y plantillas de documentos clínicos CDA R2, ofreciendo un camino para el avance de la historia clínica electrónica.⁸ Los documentos clínicos CDA, al incorporar un esquema XML que aporta un carácter semántico a los elementos del RIM y a los vocabularios utilizados, son procesables por medios informáticos, sin perder al mismo tiempo la legibilidad del documento original por parte de los usuarios. Un mismo documento, puede ser visualizado de manera transparente para el usuario, a través de navegadores o a través de dispositivos de movilidad, con total independencia de su contenido.

El CDA fue diseñado de acuerdo a principios, que por la experiencia acumulada de los miembros de HL7, eran requeridos para dar prioridad a mejorar el cuidado de los pacientes. Soporta especialmente el intercambio de documentos legibles entre usuarios, permitiendo presentar la información de forma adecuada a usuarios con diferentes requisitos o conocimientos. Promueve la duración, almacenaje e interpretación de la información más allá de formatos o tecnologías vigentes en este momento. Por su diseño, facilita un rango amplio de procesamiento posterior al intercambio, y es fácilmente compatible con muchas aplicaciones de creación y gestión documental.

Se pueden generar de diversas formas, mediante aplicaciones clientes estándar (transcripciones, dictado), a través de transformaciones desde mensajes DICOM o desde otros documentos XML y mediante herramientas de eForms: Microsoft InfoPath, Adobe Acrobat, entre otras. El documento puede ser enviado dentro de un mensaje de HL7 y puede existir de forma independiente,¹⁰ también se puede enviar usando un "web service."⁹ como ficheros.

Un receptor de un documento CDA debe ser capaz de mostrarlo según las reglas, no es

requerido que examine e interprete todas las entradas codificadas del documento, ni que valide contra alguna plantilla determinada. El generador debe poner el contenido legalmente autenticado en bloques narrativos, más allá de las entradas codificadas. En cada implementación se pueden definir responsabilidades originales de creación y recepción con respecto a secciones o entradas obligatorias. Un CDA válido debe adherir a los requerimientos de legibilidad humana, que asegura que el contenido legalmente autenticado en origen sea correctamente mostrado al que recibe. Se debe utilizar la v3 CDA, porque:

- La prioridad es la información clínica del paciente.
- Minimiza barreras técnicas de implementación.
- Promueve la longevidad de los registros clínicos.
- Facilita políticas para establecer control de requerimientos.

Un documento CDA es un objeto definido y la información completa que pueda existir fuera de un mensaje, puede incluir texto, imágenes, sonidos y otros contenidos multimedia. Contiene una cabecera y un cuerpo.

La cabecera sigue una estructura común, fácilmente consultable, que proporciona información de contexto del documento y lo identifica únicamente, provee información acerca de la autenticación, el encuentro, paciente, autor y actores involucrados. Al seguir una estructura común, bien definida, la consulta de estos campos de forma automatizada es más fácil.

El cuerpo del documento puede contener tres niveles de implementación, el nivel más bajo implica una implementación más fácil, pero se pierden muchas de las ventajas de la arquitectura CDA. El más alto ofrece una verdadera interoperabilidad semántica, pero implica un esfuerzo más alto en la implementación y requiere una madurez en los sistemas que capturan y generan los datos de los documentos.

El nivel 1 es aquel que se transmite en el cuerpo del mensaje un bloque de datos sin ninguna estructura definida, puede ser texto, una imagen, un archivo PDF, etc.

El nivel 2 sigue una estructura XML bien definida con secciones de información identificadas, pero el contenido es texto libre.

El nivel 3 agrega a cada sección, y a cada dato dentro de esas secciones, diagnósticos, unidades de medición, medicamentos, una estructura basada en el modelo común del RIM y una codificación de vocabulario estricta, con el fin de ser procesable computacionalmente. Este nivel trae muchas ventajas, la verdadera interoperabilidad semántica que permite que

los documentos sean altamente procesables, interoperables y sin ambigüedades. La norma CDA no especifica el contenido del documento, sino simplemente la estructura y semántica necesaria para su intercambio. Sin embargo, existe una forma directa de crear normas que regulen el contenido, a partir de la generación de plantillas, que restringen la norma CDA particularizando la especificación genérica para un determinado contenido.[36]

4. Machine Learning y Deep Learning en Medicina

4.1. Machine Learning

Machine Learning (ML) es un campo de la inteligencia artificial que tiene como objetivo la creación de sistemas que posean la habilidad de aprender y mejorar a partir de experiencias previas, sin necesidad de ser programados explícitamente. Este proceso de aprendizaje se basa en la observación de una gran cantidad de datos, instrucciones o experiencias previas de manera que se puedan localizar patrones y similitudes que puedan ser utilizadas en casos futuros.

Problemas y Limitaciones en Medicina

Es un hecho que áreas como la programación, la IA y la práctica clínica se unirán (y ya lo están haciendo) de forma inseparable.

Teniendo en cuenta esto, debemos plantearnos cuáles son los problemas que estamos enfrentando hoy en día, problemas que conocemos y son reales:

Los pacientes tienen enfermedades crónicas de alta prevalencia (y con complicaciones prevenibles), esto quiere decir que, afecta a un porcentaje relevante de la población y es de conocimiento abierto que al tener esas enfermedades, próximamente presentarán complicaciones. Un ejemplo de esto es el caso de la Hipertensión arterial, los pacientes desarrollarán enfermedades crónicas y; en el caso de la diabetes, los pacientes desarrollarán retinopatía diabética, si es que no se interviene a tiempo sobre ellos. [37]

Estrategias de Enfrentamiento y Procesos

Cuando existe un problema, se plantea una estrategia de enfrentamiento. Una estrategia razonable podría ser aplicar protocolos de control periódicos para mantener la enfermedad a “raya” o regulada. Las estrategias deben ser implementadas en procesos, que en cierto caso tiene que ver con:

- Ajustes frecuentes de la terapia (control).

- Exámenes de laboratorio periódicos.

Sin embargo, cuando tenemos estos procesos, nos encontramos nuevamente con problemas en salud:

- Insuficiente capital humano capacitado para poder interpretar estos exámenes, por lo que además,
- Se generan crecientes listas de espera para llegar a controlarse con un especialista.

Nuevamente se podría pensar en estrategias para solucionarlo, pero nos damos cuenta de que esto es extremadamente cíclico. Entonces, ¿qué se puede hacer al respecto?

Debemos tener en cuenta que a día de hoy la tecnología se encuentra al servicio de los problemas, por lo que podemos contar con ella para buscar estos problemas de nuestro día a día y ver cómo solucionarlos con la tecnología disponible. [38]

¿Por qué Necesitamos IA en Medicina?

Actualmente se han logrado solucionar problemas muy importantes, por ejemplo, existen grupos que han logrado obtener muy buena capacidad de diagnóstico en lesiones cutáneas utilizando Redes convolucionales. Otros grupos han trabajado en el problema de las imágenes, como radiografía de Tórax (Figura 4.1) y la detección de Neumonía.

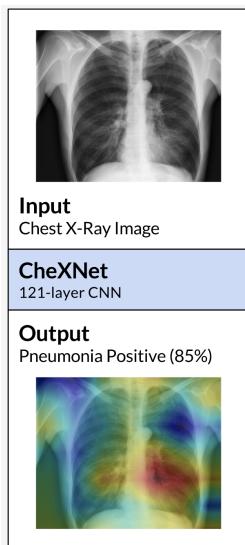


Figura 4.1: Radiografía de tórax



Figura 4.2: Radiografía de tórax de un paciente con neumonía por COVID-19, radiografía original (izquierda) y resultado con IA para neumonía (derecha) (Fotografía cortesía de UC San Diego Health)

Y otros han estudiado otros procesos de la medicina, como es la patología, particularmente la patología digital, en la cual un médico patólogo debe sentarse por horas y horas a analizar imágenes, cuando en lugar de eso, podría etiquetarlas y facilitar su propio proceso.

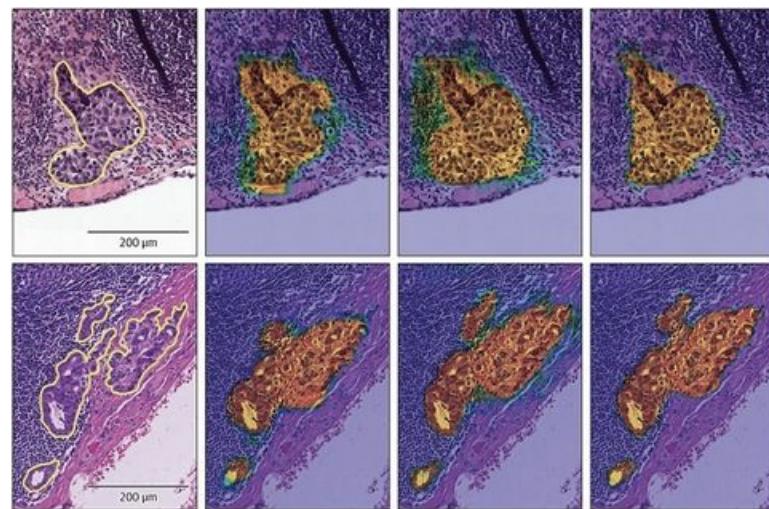


Figura 4.3: imágenes de secciones de tejido de ganglios linfáticos utilizadas para evaluar la capacidad de los algoritmos de aprendizaje profundo para detectar metástasis de cáncer. Imagen cortesía de JAMA

También hemos visto que hay avances en el diagnóstico de señales biológicas, el cual es muy interesante, ya que es uno de los pocos que tiene una manifestación concreta en la realidad, como es el caso del Apple Watch (Figura 4.5), el cual contiene la capacidad de tomar un electrocardiograma de una derivación y aplicar un algoritmo sobre esa señal para determinar si el paciente tiene o no una arritmia.

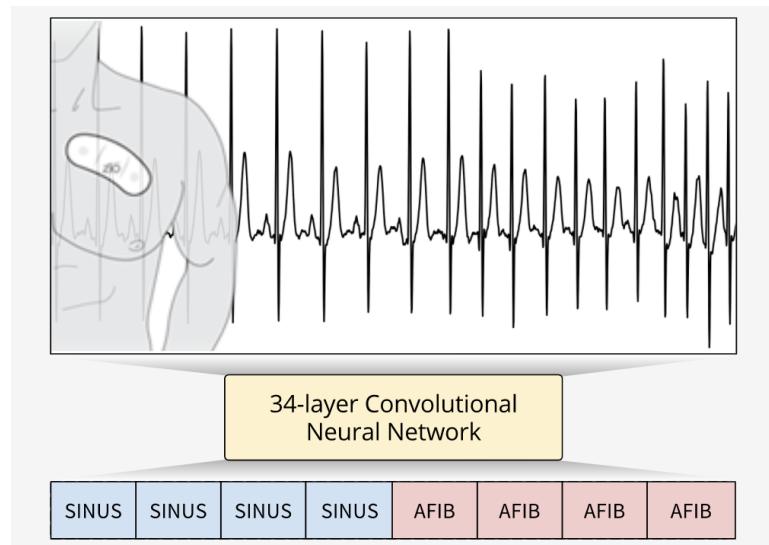


Figura 4.4: Red neuronal convolucional de 34 capas (CNN) para detectar arritmias en series de tiempo de ECG de longitud arbitraria



Figura 4.5: Detección de ritmo cardiaco por Apple Watch

Y por último, otro avance ha sido la detección de retinopatía diabética (Figura 4.5).

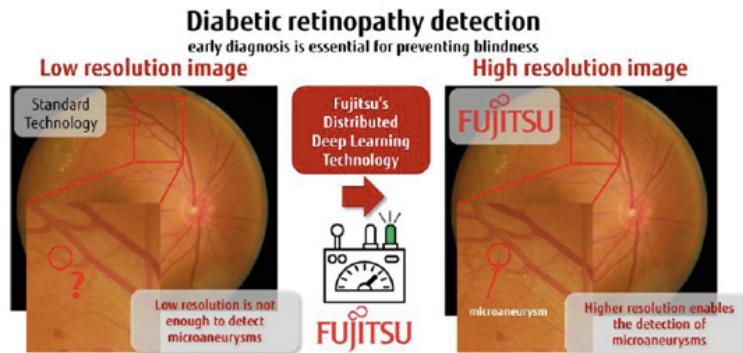


Figura 4.6: Detección temprana de retinopatía diabética. Imagen por Fujitsu

¿Por qué Hoy y No Antes?

Si bien esta tecnología existía hace 30 años, no se contaba con los requisitos fundamentales en ese entonces:

- Más datos: datos disponibles y computables.
- Mayor poder de procesamiento (GPUs/TPUs): el cual es accesible mediante internet. Cualquiera de nosotros podría levantar una instancia en alguno de los vendedores Cloud, y echar a correr un modelo.
- Reconocimiento del potencial (i.e: Alexnet): hemos reconocido la importancia y el poder de estos algoritmos, particularmente de ML.
- Democratización del conocimiento: actualmente cualquier persona podría acceder y estudiar esta tecnología si así lo quisiese.

4.2. Deep Learning

Introducción

La aparición de la inteligencia artificial (IA) en las imágenes médicas ha anunciado la tecnología disruptiva más grande desde los primeros días de Roentgen, Becquerel y Curie, pero ciertamente desde los días de Hounsfield y Anger.

La imagen médica ha dejado de ser un producto final para pasar a ser un producto intermedio que puede procesarse mediante técnicas computacionales y de ciencia de datos, como los modelos de Inteligencia Artificial (IA), lo que permite su evaluación objetiva y la extracción de parámetros cuantitativos, denominados biomarcadores de imagen, que

miden voxel a voxel (píxel en 3D) el estado de los órganos y tejidos.

El motivo por el que este cambio ha supuesto una disrupción positiva reside en las **Redes Neuronales Convolucionales** (*Convolutional Neural Networks, CNN*).

Las CNN están especialmente adaptadas para el manejo de datos bidimensionales y han tenido también un gran impacto en otros dominios, como es el caso del reconocimiento de voz. Las arquitecturas de estas redes neuronales convolucionales pueden incluir un elevado número de capas ocultas entre la entrada y la salida, que hacen que a esta tecnología se la conozca con el nombre de aprendizaje profundo o deep learning.

El **Deep Learning** se ha aplicado principalmente en radiología para la detección de lesiones y para la clasificación de las imágenes en patologías y sus subgrupos, así como para la segmentación automática de órganos. Los primeros avances fueron posibles en parte gracias a la transferencia de conocimiento (transfer learning) desde redes ya pre entrenadas para el reconocimiento y análisis de imágenes de la vida cotidiana que se han adaptado a la evaluación de imágenes radiológicas. En la actualidad, se están ya entrenando redes de manera exclusiva con datos radiológicos, lo que permite una mejora significativa en rendimiento.

Arquitectura Deep Learning

La mayoría de métodos de *machine learning* emplean arquitecturas neuronales, por lo que las arquitecturas de deep learning emplean modelización de redes neuronales, tales como:

- Deep Neural Network (DNN) o Redes Neuronales Profundas.
- Convolutional Neural Network (CNN) ó Redes Neuronales Profundas Convolucionales.
- Deep Belief Network (DBN) o Redes de Creencia Profundas.

Un Repaso por Redes Neuronales Artificiales (ANN)

Las **Redes Neuronales** son un instrumento que comienza a utilizarse en la práctica médica con el fin de ayudar al clínico en la toma de decisiones. Tienen utilidad para pronosticar la evolución de distintas patologías o intervenciones terapéuticas, clasificar o reconocer muestras anatomopatológicas, pruebas de imagen o establecer probabilidades diagnósticas a partir de síntomas o pruebas complementarias. [39]

Estas redes (ANN) son modelos matemáticos que pueden ser entrenados para aprender relaciones no lineales entre un conjunto de datos de entrada y un conjunto de datos de salida.

A nivel esquemático, una neurona artificial se representa del siguiente modo:

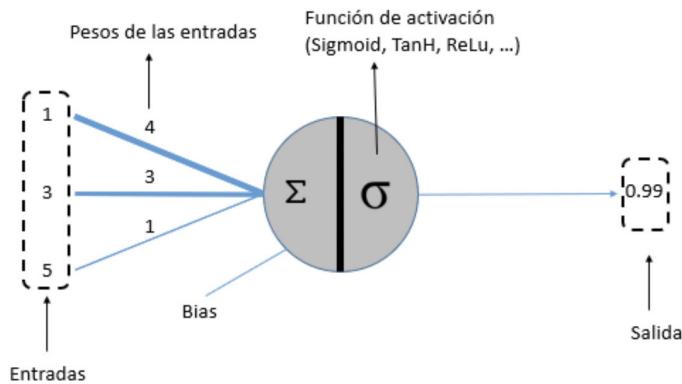


Figura 4.7: Esquema de una ANN

En el caso de las neuronas artificiales, la suma de las entradas multiplicadas por sus pesos asociados determina el “*impulso nervioso*” que recibe la neurona. Este valor, se procesa en el interior de la célula mediante una función de activación que devuelve un valor que se envía como salida de la neurona.

Del mismo modo que nuestro cerebro está compuesto por neuronas interconectadas entre sí, una red neuronal artificial está formada por neuronas artificiales conectadas entre sí y agrupadas en diferentes niveles que denominamos capas:

“*Una capa es un conjunto de neuronas cuyas entradas provienen de una capa anterior (o de los datos de entrada en el caso de la primera capa) y cuyas salidas son la entrada de una capa posterior.*”

En la siguiente imagen podemos ver una red con cuatro capas:

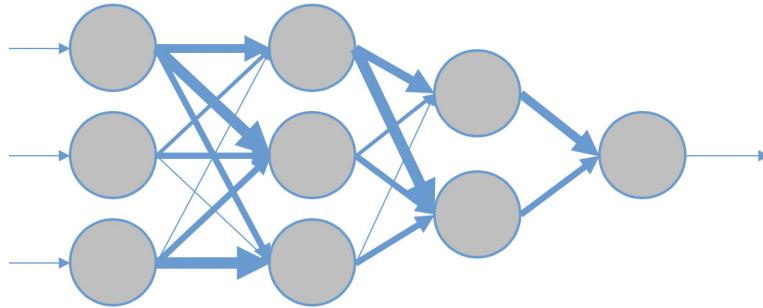


Figura 4.8: Red neuronal de cuatro capas

Las neuronas de la primera capa reciben como entrada los datos reales que alimentan a la red neuronal. Es por eso por lo que la primera capa se conoce como capa de entrada. La salida de la última capa es el resultado visible de la red, por lo que la última capa se conoce como la capa de salida. Las capas que se sitúan entre la capa de entrada y la capa de salida se conocen como capas ocultas ya que desconocemos tanto los valores de entrada como los de salida.

Una red neuronal, por lo tanto, siempre está compuesta por una capa de entrada, una capa de salida (si solo hay una capa en la red neuronal, la capa de entrada coincide con la capa de salida) y puede contener 0 o más capas ocultas. El concepto de Deep Learning nace a raíz de utilizar un gran número de capas ocultas en las redes.

En medicina la aplicación más común de estos modelos, es la clasificación de patrones con el propósito de apoyar al médico en el diagnóstico y tratamiento del paciente.

Redes Neuronales Profundas

Una **Red Neuronal Profunda (DNN)** es una Red Neuronal Artificial (ANN) con múltiples capas ocultas entre las capas de entrada y salida. De forma similar a las Redes Neuronales superficiales, la DNN pueden modelar relaciones complejas no lineales.

El objetivo principal de una red neuronal es recibir un conjunto de entradas, realizar cálculos progresivamente complejos sobre ellas y dar salida para resolver problemas del mundo real como la clasificación. Nos restringimos a alimentar redes neuronales avanzadas.

Tenemos una entrada, una salida y un flujo de datos secuenciales en una red profunda.

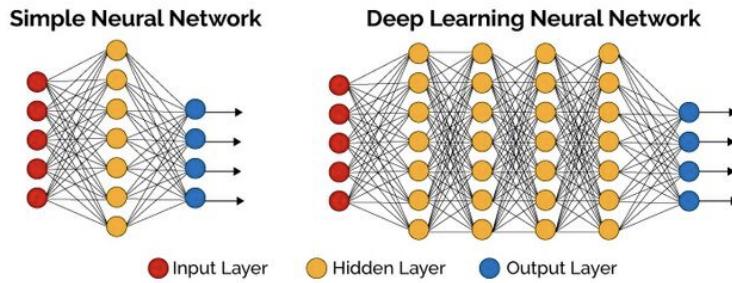


Figura 4.9: Diferencias entre una red neuronal simple y una profunda. En rojo, los inputs. En amarillo, las capas ocultas y en azul, los outputs o salidas. Imagen por MEDIUM Fávico Vázquez

Las redes neuronales se utilizan ampliamente en aprendizaje supervisado y problemas de aprendizaje de refuerzo. Estas redes se basan en un conjunto de capas conectadas entre sí.

En deep learning, la cantidad de capas ocultas, en su mayoría no lineales, puede ser grande; de alrededor de 1000 capas.

Los modelos de deep learning producen resultados mucho mejores que las redes de Machine Learning normales.

Utilizamos principalmente el método de descenso de gradiente para optimizar la red y minimizar la función de pérdida.

El entrenamiento de los conjuntos de datos forma una parte importante de los modelos de deep learning. Además, Backpropagation es el algoritmo principal en el entrenamiento de estos modelos.

El aprendizaje profundo se ocupa de la formación de redes neuronales grandes con transformaciones de salida de entrada complejas.

4.3. Redes Neuronales Profundas Convolucionales CNNs

La **Red Neuronal Convolucional (CNN)** es un tipo de Red Neuronal Artificial con aprendizaje supervisado que procesa sus capas imitando al córtex visual del ojo humano para identificar distintas características en las entradas que en definitiva hacen que pueda

identificar objetos y “ver”. Para ello, la CNN contiene varias capas ocultas especializadas y con una jerarquía: esto quiere decir que las primeras capas pueden detectar líneas, curvas y se van especializando hasta llegar a capas más profundas que reconocen formas complejas como un rostro o la silueta de un animal.

Este tipo de redes son muy efectivas para tareas de:

- Detección y categorización de objetos.
- Clasificación y segmentación de imágenes.

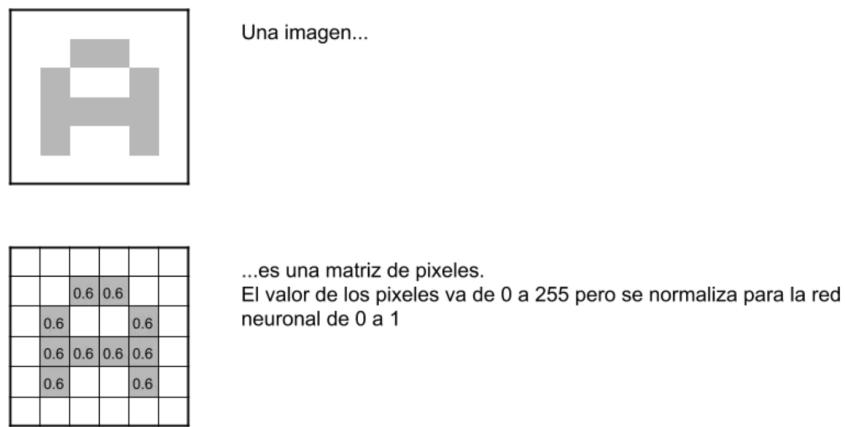


Figura 4.10

Arquitectura de una CNN

Una red neuronal convolucional es una red multicapa que consta de capas convolucionales y de reducción alternadas, y al finalmente tiene capas de conexión total como una red perceptrón multicapa.

- **Entrada:** Serán los píxeles de la imagen. Serán alto, ancho y profundidad será 1 sólo color o 3 para Red,Green,Blue.
- **Capa De Convolución:** procesará la salida de neuronas que están conectadas en «regiones locales» de entrada (es decir píxeles cercanos), calculando el producto escalar entre sus pesos (valor de píxel) y una pequeña región a la que están conectados en el volumen de entrada. Aquí usaremos por ejemplo 32 filtros o la cantidad que decidamos y ese será el volumen de salida.
- **Capa ReLu:** aplicará la función de activación en los elementos de la matriz.

- **Muestreo ó Subsampling:** Hará una reducción en las dimensiones alto y ancho, pero se mantiene la profundidad.
- **Capa tradicional:** red de neuronas feedforward que conectará con la última capa de subsampling y finalizará con la cantidad de neuronas que queremos clasificar.

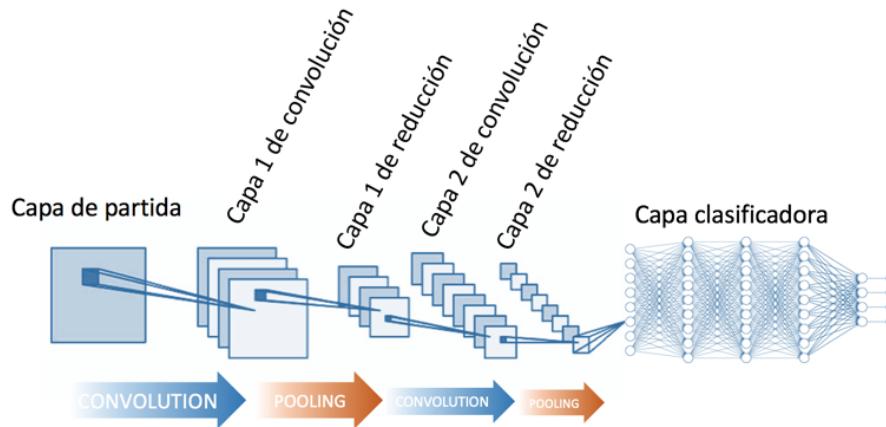


Figura 4.11: Arquitectura de red neuronal convolucional

Píxeles y Neuronas

La red toma como entrada los píxeles de una imagen. Si se tiene una imagen con apenas 28x28 píxeles de alto y ancho, eso equivale a 784 neuronas. Y eso es solo en caso de 1 color (escala de grises). Si tuviéramos una imagen a color, serían necesarios 3 canales (red, green, blue) y entonces usaríamos $28 \times 28 \times 3 = 2352$ neuronas de entrada. Esa es nuestra capa de entrada.

Pre-procesamiento

Antes de alimentar la red, es bueno recordar que como entrada conviene normalizar los valores. Los colores de los píxeles tienen valores que van de 0 a 255, se realizará una transformación de cada pixel: «valor/255» y quedará siempre un valor entre 0 y 1.

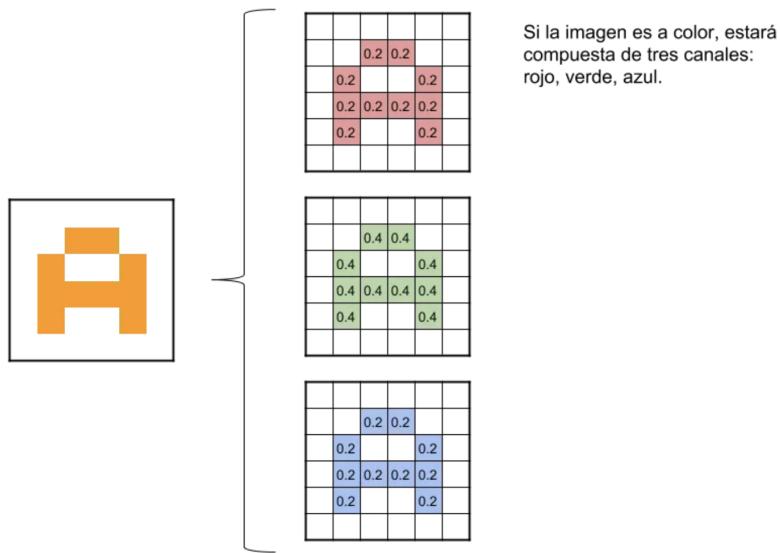


Figura 4.12: Canales de una imagen a color

Convoluciones

Ahora es cuando comienza el «*procesado distintivo*» de las CNN. Es decir, donde se hacen las llamadas «*convoluciones*»: Estas consisten en tomar grupos de píxeles cercanos de la imagen de entrada e ir operando matemáticamente (producto escalar) contra una pequeña matriz que se llama kernel. Ese kernel, por ejemplo, de tamaño 3x3 pixels recorre todas las neuronas de entrada (de izquierda-derecha, de arriba-abajo) y genera una nueva matriz de salida, que en definitiva será la nueva capa de neuronas ocultas.

NOTA: Si la imagen fuera a color, el kernel realmente sería de 3x3x3, es decir: un filtro con 3 kernels de 3x3; luego esos 3 filtros se suman (y se le suma una unidad bias) y conformarán 1 salida (como si fuera 1 solo canal).

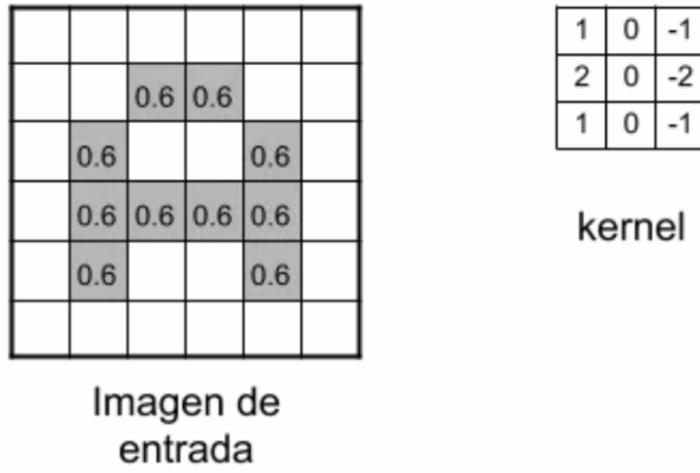


Figura 4.13: El kernel tomará inicialmente valores aleatorios (1) y se irán ajustando mediante Backpropagation. Una mejora es hacer que siga una distribución normal siguiendo simetrías, pero sus valores son aleatorios

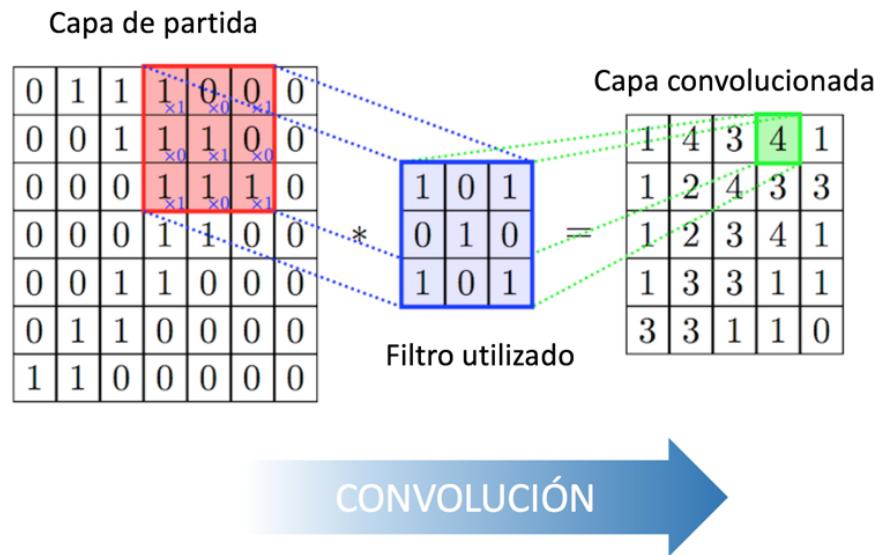


Figura 4.14: Convolución

Después de aplicar la convolución se le aplica a los mapas de características una función

de activación.

La función de activación que suele ser recomendada para este tipo de red es **sigmoide** **ReLU**, seleccionando una tasa de aprendizaje adecuada y monitorizar la fracción de neuronas muertas, también se puede probar con **Leaky ReLU** o Maxout, pero nunca utilizar sigmoide logística.

Filtro: Conjunto de Kernels

En realidad, no se aplica 1 solo kernel, si no muchos kernels (su conjunto se llama filtros). Por ejemplo, en esta primer convolución se podría tener 32 filtros, con lo cual, realmente se obtiene 32 matrices de salida (este conjunto se conoce como «feature mapping»), cada una de 28x28x1 dando un total del 25.088 neuronas para la **primera capa oculta de neuronas**.

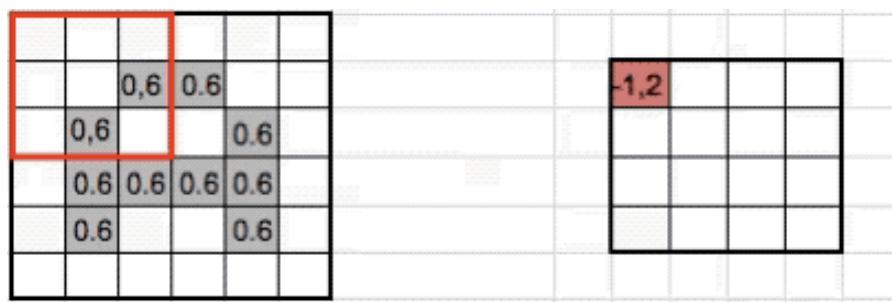


Figura 4.15: Kernel realizando el producto matricial con la imagen de entrada y desplazando de a 1 pixel de izquierda a derecha y de arriba-abajo y va generando una nueva matriz que compone al mapa de features

A medida que se va desplazando el kernel se irá obteniendo una «*nueva imagen*» filtrada por el kernel. En esta primer convolución y siguiendo con el ejemplo anterior, es como si se obtuvieran 32 «*imágenes filtradas nuevas*». Estas imágenes nuevas lo que están “dibujando” son ciertas características de la imagen original. Esto ayudará en el futuro a poder distinguir un objeto de otro (por ej. gato o perro).

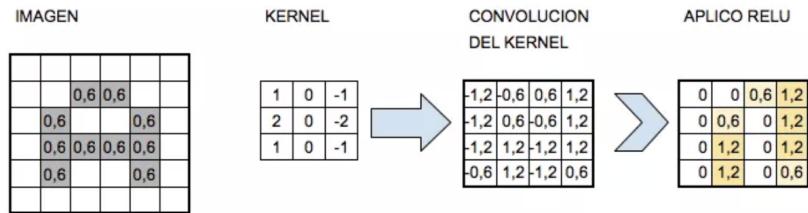


Figura 4.16: La imagen realiza una convolución con un kernel y aplica la función de activación, en este caso ReLu. Y es cuando finalmente se obtiene un mapa de detección de características

Función de Activación

La función de activación se encarga de devolver una salida a partir de un valor de entrada, normalmente el conjunto de valores de salida en un rango determinado como (0,1) o (-1,1).

Se buscan funciones que las derivadas sean simples, para minimizar con ello el coste computacional. [40, 41]

Tipos de Función de Activación

Sigmoid – Sigmoide

La función sigmoide transforma los valores introducidos a una escala (0,1), donde los valores altos tienden de manera asintótica a 1 y los valores muy bajos tienden de manera asintótica a 0.

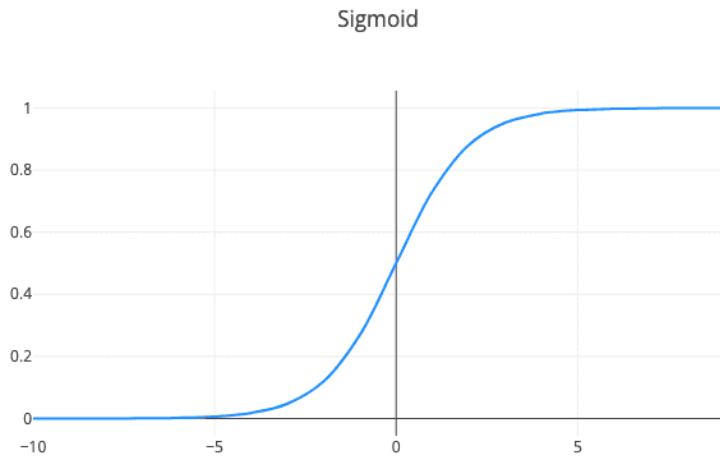


Figura 4.17: Función de activación Sísmoide

$$f(x) = \frac{1}{1 + e^{-x}}$$

Características de la función sigmoide:

- Satura y mata el gradiente.
- Lenta convergencia.
- No está centrada en el cero.
- Está acotada entre 0 y 1.
- Buen rendimiento en la última capa.

Tanh – Tangent Hyperbolic – Tangente Hiperbólica

La función tangente hiperbólica transforma los valores introducidos a una escala (-1,1), donde los valores altos tienden de manera asintótica a 1 y los valores muy bajos tienden de manera asintótica a -1.

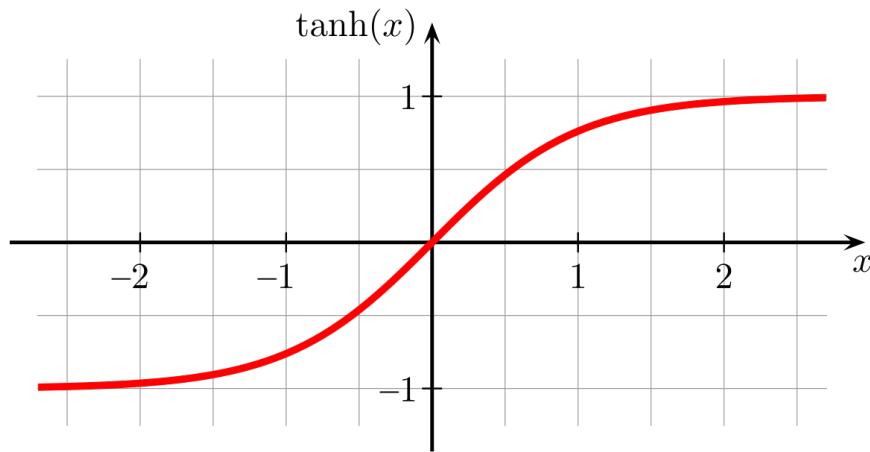


Figura 4.18: Función de activación Tangente Hiperbólica

$$f(x) = \frac{2}{1 + e^{-2x}} - 1$$

Características de la función tangente hiperbólica:

- Muy similar a la sigmoide.
- Satura y mata el gradiente.
- Lenta convergencia.
- Centrada en 0.
- Está acotada entre -1 y 1.
- Se utiliza para decidir entre una opción y la contraria.
- Buen desempeño en redes recurrentes.

ReLU – Rectified Linear Unit

La función ReLU transforma los valores introducidos anulando los valores negativos y dejando los positivos tal y como entran.

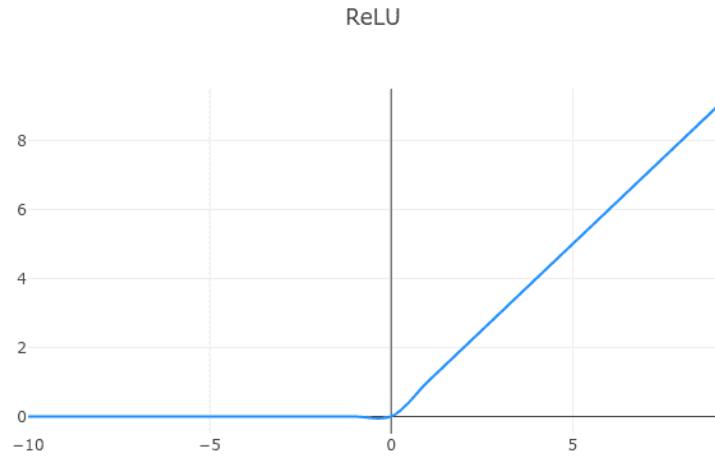


Figura 4.19: Función de activación ReLU

$$f(x) = \max(0, x) = \begin{cases} 0 & \text{for } x < 0 \\ x & \text{for } x \geq 0 \end{cases}$$

Características de la función ReLU:

- Activación Sparse – solo se activa si son positivos.
- No está acotada.
- Se pueden morir demasiadas neuronas.
- Se comporta bien con imágenes.
- Buen desempeño en redes convolucionales.

Leaky ReLU – Rectified Linear Unit

La función Leaky ReLU transforma los valores introducidos multiplicando los negativos por un coeficiente rectificativo y dejando los positivos seg\xf1n entran.

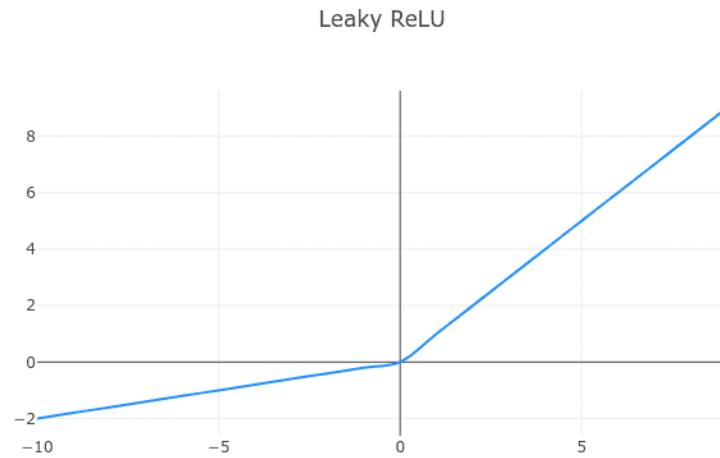


Figura 4.20: Función de activación Leaky ReLU

$$f(x) = \begin{cases} 0 & \text{for } x < 0 \\ a \cdot x & \text{for } x \geq 0 \end{cases}$$

Características de la función Leaky ReLU:

- Similar a la función ReLU.
- Penaliza los negativos mediante un coeficiente rectificador.
- No está acotada.
- Se comporta bien con imágenes.
- Buen desempeño en redes convolucionales.

Softmax – Rectified Linear Unit

La función Softmax transforma las salidas a una representación en forma de probabilidades, de tal manera que el sumatorio de todas las probabilidades de las salidas de 1.

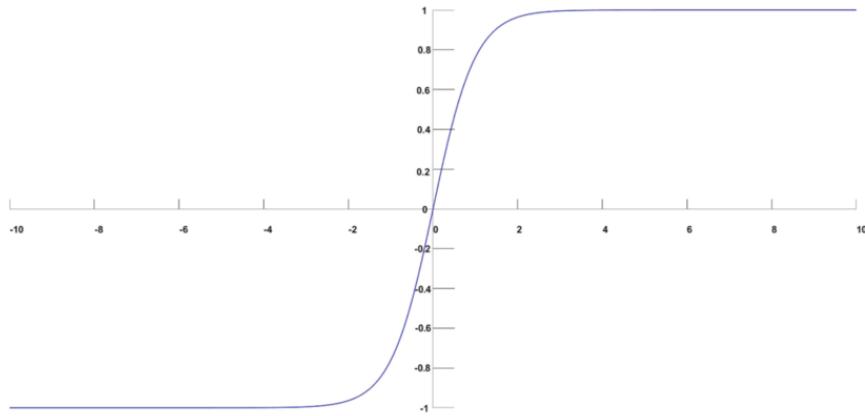


Figura 4.21: Función de activación Softmax

$$f(Z)_j = \frac{e^{z_j}}{\sum_{k=1}^K e^{z_k}}$$

Características de la función Softmax:

- Se utiliza cuando queremos tener una representación en forma de probabilidades.
- Está acotada entre 0 y 1.
- Muy diferenciable.
- Se utiliza para normalizar tipo multiclas.
- Buen rendimiento en las últimas capas.

Las CNNs y el Modelo Matemático de Backpropagation de las Redes Neuronales

El proceso es similar al de las redes tradicionales en las que tenemos una entrada y una salida esperada (por eso le llamamos aprendizaje supervisado) y mediante ese procesamiento de ajuste hacia adelante y hacia atrás, vamos mejorando el valor de los pesos de las interconexiones entre capas de neuronas y a medida que iteramos esos pesos se ajustan hasta ser óptimos.

En el caso de la CNN, deberemos ajustar el valor de los pesos de los distintos kernels. Esto es una gran ventaja al momento del proceso de aprendizaje pues como vimos cada kernel es de un tamaño reducido, en nuestro ejemplo en la primer convolución es de tamaño de 3x3, eso son sólo 9 parámetros que debemos ajustar en 32 filtros dan un total de 288 parámetros. En comparación con los pesos entre dos capas de neuronas «tradicionales»:

una de 748 y otra de 6272 en donde están TODAS interconectararlas con TODAS y eso equivaldría a tener que entrenar y ajustar más de 4,5 millones de pesos (sólo para 1 capa).

4.4. Red Convolucional U-Net

A pesar de que las redes convolucionales han existido durante mucho tiempo , su uso estaba limitado debido a problemas computacionales (redes de gran tamaño, dataset pesados, etc).

El uso de las redes profundas se inició gracias a Krizhevsky.

Krizhevsky fué capaz de realizar un entrenamiento supervisado de una red grande con 8 capas y millones de parámetros en el conjunto de datos ImageNet con 1 millón de imágenes de entrenamiento.

Desde entonces, incluso las redes más grandes y profundas han podido ser entrenadas.

El uso típico de las redes convolucionales es en realizar predicciones de clasificación, donde la salida a una imagen es una etiqueta de clase única.

Sin embargo, en muchas predicciones en las cuales los datos de entrada son imágenes, especialmente en el procesamiento de imágenes biomédicas, la salida deseada debe incluir la localización, es decir, se supone que se debe asignar una etiqueta de clase a cada píxel.

Además, miles de imágenes de entrenamiento están generalmente fuera del alcance en proyectos biomédicos.

Ciresan entrenó una red en una configuración de ventana deslizante para predecir la etiqueta de clase de cada píxel proporcionando una región local (parche) alrededor de ese píxel como entrada.

Primero, dicha red puede localizar. En segundo lugar, los datos de entrenamiento (cada píxel tiene su etiqueta) son mucho mayores que el número de imágenes de entrenamiento.

La red creada por Ciresan et al presenta los siguientes problemas:

- Es una red de alto coste computacional (es lenta): la red debe ejecutarse por separado para cada parche, y hay mucha redundancia debido a parches superpuestos.

- Existe un desempate entre la precisión de localización y el uso del entorno: los parches más grandes requieren más capas de agrupación máxima que reducen la precisión de localización, mientras que los parches pequeños permiten que la red vea solo un pequeño entorno.

Para solventar dichos problemas, enfoques más recientes han propuesto un resultado clasificador que tiene en cuenta las características de varias capas, permitiendo así una buena localización y uso del entorno al mismo tiempo.

La buena localización y el uso del contexto son posibles al mismo tiempo.

La red U-Net [42] presenta una arquitectura más elegante, la llamada red totalmente convolucional es capaz de funcionar con muy pocas imágenes de entrenamiento y producir segmentaciones más precisas (Figura 4.22).

La idea principal es complementar una red convolucional habitual con capas sucesivas, donde los operadores de agrupación son reemplazados por operadores de muestreo ascendente aumentando así la resolución de la salida.

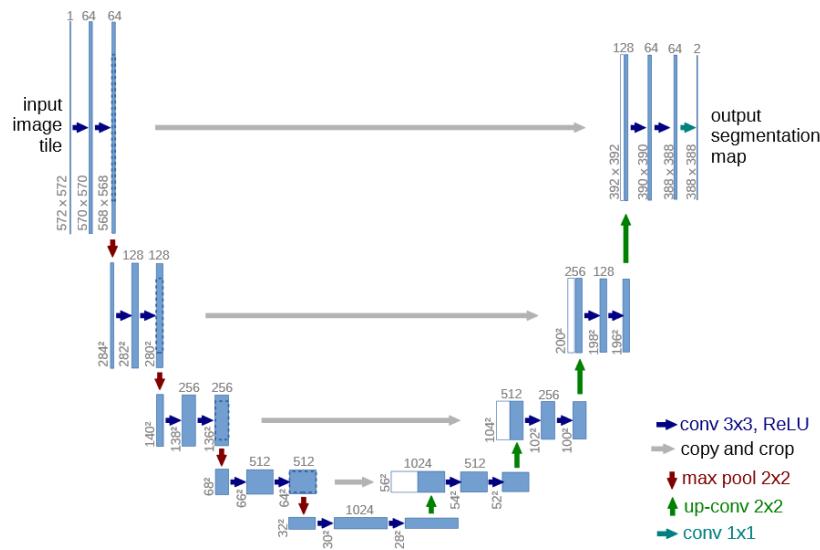


Figura 4.22: Arquitectura U-Net

En la parte de muestreo superior existe una gran cantidad de canales de características permitiendo que la red propague información del entorno a capas de mayor resolución.

Como consecuencia, el camino expansivo es más o menos simétrico al camino de contratación, produciendo una arquitectura en forma de U.

Las capas no están completamente conectadas y solo usa la parte válida de cada convolución, es decir, el mapa de segmentación sólo contiene los píxeles, para los cuales el entorno completo está disponible en la imagen de entrada.

Esta estrategia permite la segmentación sin problemas de imágenes grandes por estrategia de superposición de mosaicos (ver Figura 4.23).

Para predecir los píxeles en la región del borde de la imagen, el entorno que falta se extrae reflejando la imagen de entrada. Esta estrategia de mosaico es importante para aplicar la red a imágenes grandes, ya que de lo contrario la resolución estaría limitada por la memoria de la GPU.

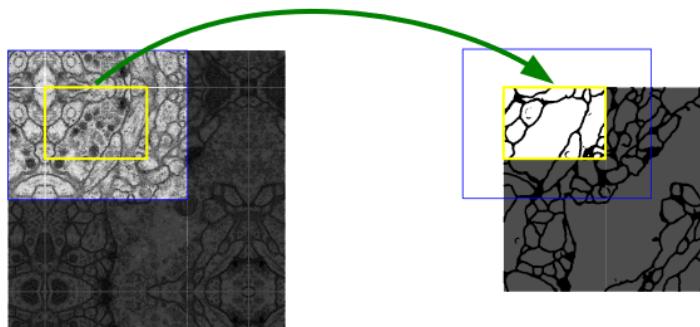


Figura 4.23: Superposición de mosaicos. La predicción de la segmentación en el área amarilla requiere datos de imagen dentro del área azul como entrada. Los datos de entrada faltantes se extrapolan por duplicación

Arquitectura U-Net

La arquitectura de la red se muestra en la Figura 4.22.

Como podemos observar, existen dos tipos de rutas (una por el lado izquierdo y otra por el lado derecho). La arquitectura situada al lado izquierdo es la típica que podemos encontrar en una red convolucional. Consiste en la aplicación repetitiva de dos capas convolucionales 3x3 cada una activadas mediante la función Relu y un kernel de 2x2 con la finalidad de disminuir la resolución.

En cada paso de reducción del tamaño de muestra, se duplica el número de canales.

La arquitectura situada en el lado derecho consiste en una entrada de muestra de manera ascendente seguido de una convolución 2x2 reduciendo a la mitad el número de canales funcionales, una concatenación con las capas situadas a la izquierda y dos convoluciones de 3x3 con una activación Relu.

En la capa final, se utiliza una convolución 1x1.

En total la red consta de 23 capas convolucionales.

Entrenamiento

Las imágenes de entrada y sus correspondientes mapas de segmentación se utilizan para entrenar la red usando el descenso de gradiente estocástico de Caffe.

Debido a circunvoluciones sin relleno, la imagen de salida presenta un menor tamaño que la imagen de entrada.

Para minimizar los gastos generales en el entrenamiento de la red así como el máximo aprovechamiento de la memoria de la GPU, se recomienda tomar grandes mosaicos como datos de entrada sobre un gran tamaño de lote y usar un gran impulso (0.99), permitiendo así que una gran cantidad de las muestras de entrenamiento vistas previamente por la red determinen la actualización de los pesos de las neuronas en el paso de optimización actual.

La función de energía se calcula mediante la función soft-max combinado con la función de pérdida de entropía cruzada.

- **Función softmax:** Se usa para “comprimir” un vector K-dimensional ,z, de valores arbitrarios en un vector K-dimensional, $\sigma(z)$, de valores reales en el rango [0,1]. Su función es la siguiente:

$$p_k(X) = \exp(a_k(X)) / (\sum_{k'=1}^K \exp(a_{k'}(X)))$$

- **Entropía cruzada:** Métrica que puede utilizarse para reflejar la precisión de los pronósticos probabilísticos y está estrechamente vinculada con la estimación por máxima verosimilitud.

Para dos variables aleatorias discretas p y q, la entropía cruzada se define:

$$H(p, q) = - \sum_x p(x) \log q(x)$$

En la práctica, debido a que p no se conoce, la entropía cruzada se calcula empíricamente a partir de las observaciones, suponiendo que son igualmente probables. De esta forma, se obtiene la siguiente fórmula:

$$H(q) = -\frac{1}{N} \sum_x \log q(x)$$

4.5. Autoencoders

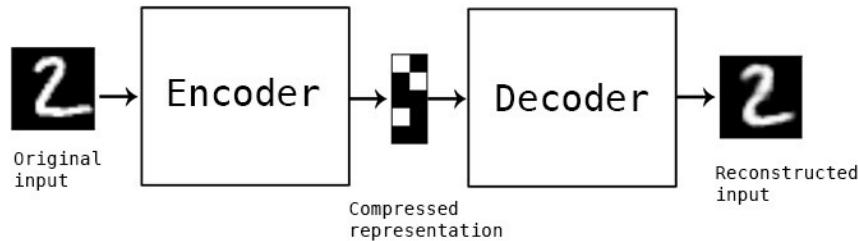


Figura 4.24

Los **autoencoders**[43], son redes neuronales cuyos valores de entrada son iguales a los de la salida (Figura 4.25), por lo que emplea aprendizaje no supervisado.

Tienen como objetivo generar nuevos datos, primero comprimiendo la entrada en un espacio de variables latentes (cuando coge la entrada, crea un espacio con variables “inferidas”), y luego reconstruye la salida en base a la información adquirida.

El **autoencoder**, por diseño, reduce las dimensiones de los datos al aprender a ignorar el ruido de estos.

Típicamente los autoencoders son simétricos con respecto a la capa central y se componen de 2 partes principales: Desde la capa de entrada hasta la capa central, ambas incluidas, se denomina *encoder*(Figura (a) 4.26), y desde la capa central hasta la capa de salida,

decoder (Figura (b) 4.26). Esta nomenclatura se debe a que si se divide el autoencoder en dos redes neuronales, funcionan como un algoritmo de codificación (Figura 4.26): la red neuronal cuya entrada es la entrada del autoencoder y su salida la capa central, actúa como una función de compresión de información, mientras que la red cuya capa de entrada es la capa central del autoencoder y su capa de salida es la capa de salida del autoencoder, actúa como una función de descompresión. El encoder funciona también como extractor de características.

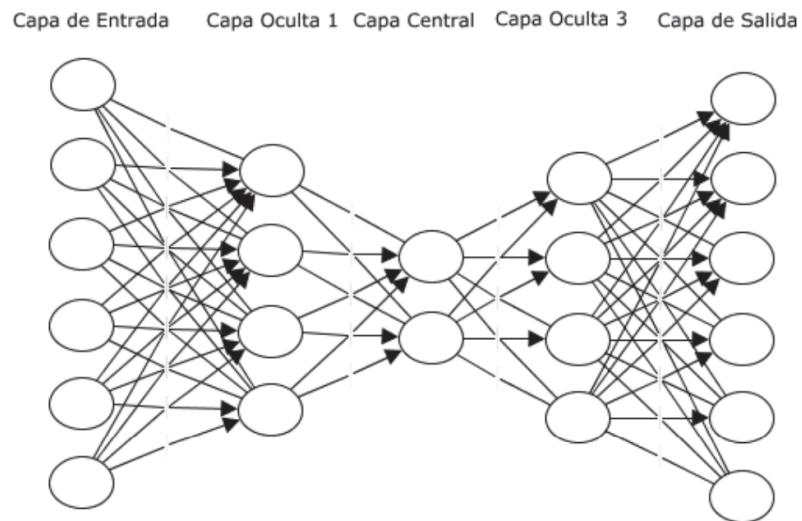


Figura 4.25: Autoencoder

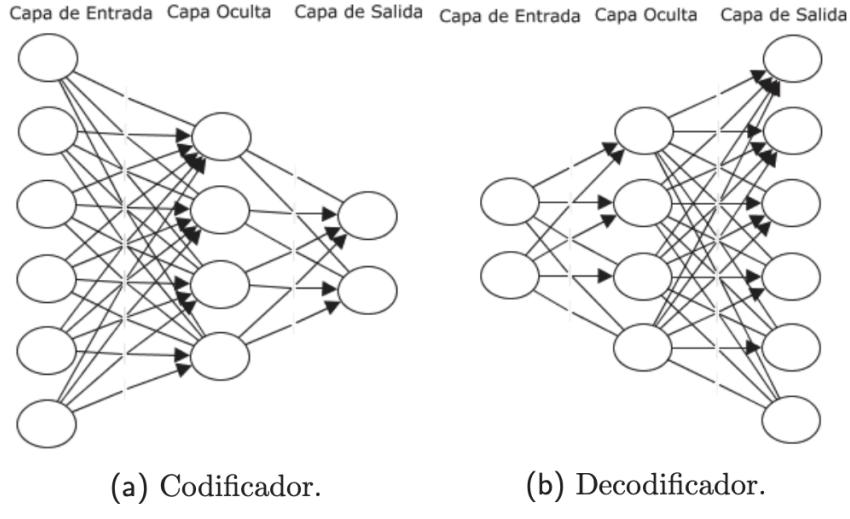


Figura 4.26: Representación de encoder y decoder

Si bien esta red se compone de 2 partes, como hemos visto anteriormente, debemos conocer los 4 conceptos fundamentales que conforman esta red:

1. **Encoder:** la parte de la red que comprime la entrada en un espacio de variables latentes que puede representarse mediante la función de codificación $h = f(x)$
2. **Bottleneck:** es la capa que contiene la representación comprimida de los datos de entrada. Estas son las dimensiones más bajas posibles de los datos de entrada.
3. **Decoder:** la parte en la que el modelo aprende a reconstruir los datos de la representación codificada para estar lo más cerca posible de la entrada original. Se representa mediante la función de decodificación: $r = g(h)$
4. **Reconstruction Loss:** este es el método que mide la precisión del rendimiento del decoder y qué tan cerca está la salida de la entrada original.

Arquitectura de un Autoencoder

El autoencoder se puede describir con la siguiente función $d(f(x)) = r$ donde r es similar a la entrada original de x .

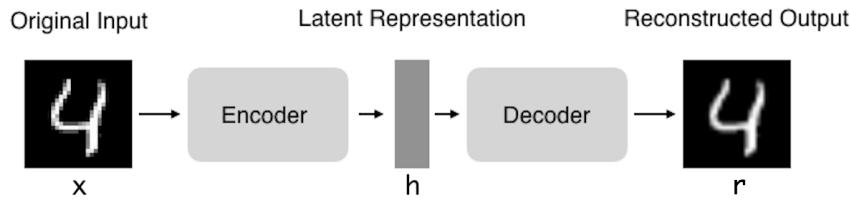


Figura 4.27: Arquitectura de un autoencoder

¿Por qué Copiar la Entrada en la Salida?

Lo que esperamos es que, cuando entrenamos un autoencoder y copiamos la entrada, es espacio de variables latentes h pueda tomar características útiles para nosotros.

Esto se puede lograr imponiendo límites a la acción de codificación, forzando el espacio h que sea de menores dimensiones que x . En este caso, el autoencoder viene llamado *undercomplete*. Al entrenar el espacio *undercomplete*, traemos el autoencoder que captura las características más relevantes de los datos de entrenamiento. Si no les damos suficientes restricciones, la red se limita a la tarea de copiar la entrada en la salida, sin individuar ninguna información útil sobre la distribución de datos. Esto puede suceder incluso cuando el tamaño de subespacio latente tiene el mismo tamaño que el espacio inicial; en el caso de cuando tenemos un caso de autoencoder *overcomplete*, significa que el tamaño de espacio de variables latentes es mayor que las entradas. En estos casos, incluso con simples encoder y decoder lineales, podemos copiar la entrada en la salida sin aprender nada sobre los datos.

Idealmente, es posible entrenar con éxito cualquier arquitectura basada en autoencoder, eligiendo adecuadamente los parámetros y la capacidad de cada encoder-decoder en función de la complejidad de los datos a modelar.

El Uso de los Autoencoders

Hasta la fecha, la reducción de ruido y la reducción de dimensionalidad para la visualización de datos se consideran las aplicaciones más interesantes de autoencoders, con el ajuste apropiado.

Con la configuración apropiada de la dimensionalidad y las restricciones relacionadas en la dispersión de los datos, a través de autoencoder es posible obtener proyecciones en subespacios de mayor interés en comparación con métodos lineales como PCA.

Los autoencoders se entranan automáticamente a través de datos de muestra. Esto significa que es fácil entrenar la red para tener un buen rendimiento en tipos similares de entrada, sin la necesidad de generalizar. Es decir, solo la compresión realizada en datos similares a los utilizados en el conjunto de entrenamiento tendrá buenos resultados, pero si se realiza en diferentes datos no será muy efectiva. Otras técnicas de compresión como JPEG serán más capaces de realizar esta función.

Estas redes neuronales están capacitadas para conservar la mayor cantidad de información posible cuando se insertan en el encoder y luego en el decoder, pero también para que las nuevas representaciones adquieran diferentes tipos de propiedades. Se describen cuatro tipos de autoencoders:

Tipos de Autoencoders

- Vanilla Autoencoder.
- Autoencoder Multicapa.
- Autoencoder Convolucional.
- Autoencoder Regularizado.

Vanilla Autoencoder

Es la forma más simple, caracterizada por una red de tres capas, es decir, una red neuronal con una sola capa oculta. La entrada y la salida son las mismas y, en nuestro caso, aprendemos cómo reconstruir la entrada mediante el uso del optimizador Adam , como función de pérdida, la desviación mínima del cuadrado.

En esta tipología estamos en el caso de encoder undercomplete donde la capa oculta, con un tamaño igual a 64, es más pequeña que la de la entrada (784), obtenida de la siguiente fórmula $28x28x1$. Esta restricción fuerza a la red neuronal a aprender de una representación de datos comprimida.

```

input_size = 784
hidden_size = 64
output_size = 784

x = Input(shape=(input_size,))

# Encoder
h = Dense(hidden_size, activation='relu')(x)

# Decoder
r = Dense(output_size, activation='sigmoid')(h)

autoencoder = Model(input=x, output=r)
autoencoder.compile(optimizer='adam', loss='mse')

```

Figura 4.28

Autoencoder Multicapa

Si solo una capa oculta no es suficiente, podemos extender el autoencoder a lo largo de la dimensión de profundidad. Nuestra implementación ahora usa tres capas ocultas para una mejor generalización, pero también tendremos que hacer que la red sea simétrica utilizando la capa intermedia.

```

input_size = 784
hidden_size = 128
code_size = 64

x = Input(shape=(input_size,))

# Encoder
hidden_1 = Dense(hidden_size, activation='relu')(x)
h = Dense(code_size, activation='relu')(hidden_1)

# Decoder
hidden_2 = Dense(hidden_size, activation='relu')(h)
r = Dense(input_size, activation='sigmoid')(hidden_2)

autoencoder = Model(input=x, output=r)
autoencoder.compile(optimizer='adam', loss='mse')

```

Figura 4.29

Autoencoder Convolucional

Surge una pregunta: ¿podrían usarse autoencoders con convoluciones en lugar de capas conectadas por completo?

Y la respuesta es sí. El principio es el mismo, pero se usan vectores tridimensionales en lugar de vectores unidimensionales. La imagen de entrada se muestrea para obtener una representación latente, es decir, una reducción dimensional, lo que obliga al autoencoder a aprender de una versión comprimida de la imagen.

```
x = Input(shape=(28, 28,1))

# Encoder
conv1_1 = Conv2D(16, (3, 3), activation='relu', padding='same')(x)
pool1 = MaxPooling2D((2, 2), padding='same')(conv1_1)
conv1_2 = Conv2D(8, (3, 3), activation='relu', padding='same')(pool1)
pool2 = MaxPooling2D((2, 2), padding='same')(conv1_2)
conv1_3 = Conv2D(8, (3, 3), activation='relu', padding='same')(pool2)
h = MaxPooling2D((2, 2), padding='same')(conv1_3)

# Decoder
conv2_1 = Conv2D(8, (3, 3), activation='relu', padding='same')(h)
up1 = UpSampling2D((2, 2))(conv2_1)
conv2_2 = Conv2D(8, (3, 3), activation='relu', padding='same')(up1)
up2 = UpSampling2D((2, 2))(conv2_2)
conv2_3 = Conv2D(16, (3, 3), activation='relu')(up2)
up3 = UpSampling2D((2, 2))(conv2_3)
r = Conv2D(1, (3, 3), activation='sigmoid', padding='same')(up3)

autoencoder = Model(input=x, output=r)
autoencoder.compile(optimizer='adam', loss='mse')
```

Figura 4.30

Autoencoder Regularizado

Existen otros métodos mediante los cuales podemos restringir la reconstrucción del autoencoder como una alternativa a la imposición de una dimensionalidad reducida de las capas ocultas. En lugar de limitar la capacidad del modelo manteniendo una arquitectura poco profunda del encoder y decoder , así como una reducción forzada, los autoencoders regularizados utilizan una función de pérdida para alentar al modelo a asumir propie-

dades que van más allá de la simple capacidad de copiar la entrada en la salida . En la práctica, encontramos dos tipos diferentes: autoencoder disperso y Denoising autoencoder.

- **Autoencoder escaso:** generalmente se utilizan para la clasificación. Al entrenar un autoencoder, las unidades ocultas en la capa intermedia se activan con demasiada frecuencia. Para evitar esto, debemos reducir su tasa de activación limitándola a una fracción de los datos de entrenamiento. Esta restricción se denomina restricción de dispersión, porque cada unidad se activa solo mediante un tipo de entrada predefinido.

```
input_size = 784
hidden_size = 64
output_size = 784

x = Input(shape=(input_size,))

# Encoder
h = Dense(hidden_size, activation='relu', activity_regularizer=regularizers.l1(10e-5))(:)

# Decoder
r = Dense(output_size, activation='sigmoid')(h)

autoencoder = Model(input=x, output=r)
autoencoder.compile(optimizer='adam', loss='mse')
```

Figura 4.31

Otro cambio que podemos hacer durante la fase de reconstrucción es agregar un término de regularización a nuestra función de pérdida: por ejemplo, un regulador de tipo 1 introduce una penalización en la función de pérdida, que actúa durante la fase de optimización. Como resultado, tendremos menos activaciones que las de un autoencoder de Vanilla.

- **Denoising autoencoder:** En lugar de agregar penalización a la función de pérdida, podemos hacer que el objeto cambie, agregando ruido a la imagen de entrada y haciendo que el autoencoder aprenda a eliminarlo de manera autónoma. Esto significa que la red extraerá solo la información más relevante y aprenderá de una representación sólida de los datos.

```

x = Input(shape=(28, 28, 1))

# Encoder
conv1_1 = Conv2D(32, (3, 3), activation='relu', padding='same')(x)
pool1 = MaxPooling2D((2, 2), padding='same')(conv1_1)
conv1_2 = Conv2D(32, (3, 3), activation='relu', padding='same')(pool1)
h = MaxPooling2D((2, 2), padding='same')(conv1_2)

# Decoder
conv2_1 = Conv2D(32, (3, 3), activation='relu', padding='same')(h)
up1 = UpSampling2D((2, 2))(conv2_1)
conv2_2 = Conv2D(32, (3, 3), activation='relu', padding='same')(up1)
up2 = UpSampling2D((2, 2))(conv2_2)
r = Conv2D(1, (3, 3), activation='sigmoid', padding='same')(up2)

autoencoder = Model(input=x, output=r)
autoencoder.compile(optimizer='adam', loss='mse')

```

Figura 4.32

4.6. Tipos de Ruido

La imagen de resonancia magnética (IRM), es una notable técnica no invasiva, que permite crear imágenes de la anatomía del cuerpo humano en cualquier proyección sin mover al paciente. La calidad visual de esta imagen desempeña un papel fundamental en la precisión del análisis clínico; sin embargo, se puede ver severamente degradada por ruido. En la actualidad aunque la tecnología del resonador magnético ha aumentado notablemente, la IRM sigue siendo afectada por el ruido, este se incorpora inevitablemente durante el proceso de adquisición. El ruido no solo afecta al diagnóstico clínico, si no también tareas como segmentación, reconocimiento y análisis computarizado automático.

El término ruido es usado habitualmente para designar señales no deseadas que tienden a perturbar transmisión y procesamiento de las señales en sistemas de comunicaciones, y sobre las cuales no se tiene control.

En la práctica existen muchas potenciales fuentes de ruido en los sistemas de comunicaciones, estas fuentes pueden ser:

- Externas al sistema por ej: ruido atmosférico, ruido galáctico, ruido hecho por el hombre.
- Interna al sistema: estos surgen de fluctuaciones espontáneas de corriente o voltaje en circuitos eléctricos. Estos representan una limitación básica en la transmisión y detección de señales : los dos ejemplos más comunes de fluctuaciones espontáneas en circuitos eléctricos son : shot noise y thermal noise (ruido de disparo y ruido térmico).

Durante este proyecto se han usado los siguientes tipos de ruido para perturbar las imágenes originales:

4.6.1. Salt and Pepper:

Este tipo de ruido [44] se caracteriza principalmente por cubrir de forma dispersa toda la imagen con una serie de píxeles blancos y negros (en el caso de las imágenes en blanco y negro) o en el caso de imagen en color, el ruido cambia el número de apariciones de combinación de colores (Figura 4.33).

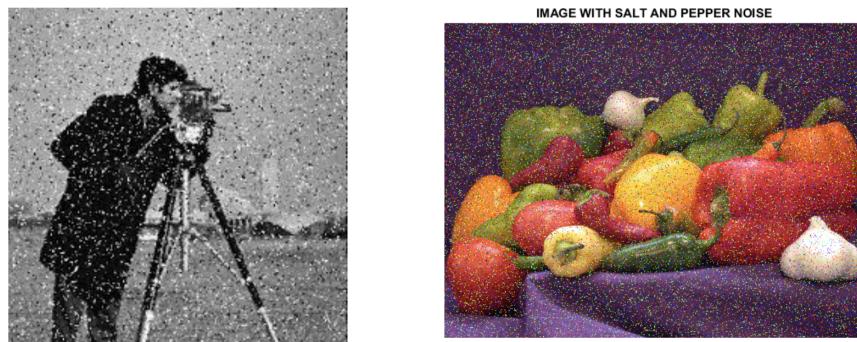


Figura 4.33: Imágenes con ruido Salt and Pepper

cambia el número de apariciones de combinación de colores (Figura 4.33).

Una forma efectiva para la reducción de este tipo de ruido es mediante el uso de filtros medianos o filtros morfológicos.

Los filtros medianos más utilizados son:

- Filtro mediano estándar (SMF).

- Filtro mediano adaptativo (AMF).
- Filtro mediano de cambio progresivo (PSMF).
- Filtro mediano de prevención de detalles (DPF).

El algoritmo de filtrado varía de un algoritmo a otro según la precisión de aproximación para el píxel ruidoso utilizando sus píxeles circundantes.

El ruido Salt and Pepper generalmente es causado por un defecto del sensor de la cámara, fallo de software o fallo del hardware en la captura o transmisión de imágenes.

En este tipo de ruido, solo una proporción de todos los píxeles de la imagen están dañados mientras que el resto de píxeles no se encuentran afectados (permanecen sin cambios).

4.6.2. Gaussiano:

El ruido Gaussiano[45] generalmente perturba los valores grises en las imágenes digitales y se caracteriza por tener una función de densidad de probabilidad igual a la de la distribución normal, también conocida como distribución gaussiana (Figura 4.34).

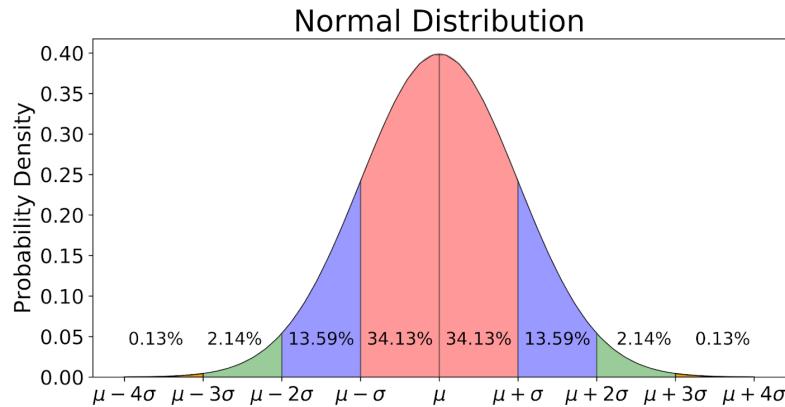


Figura 4.34: Distribución Gaussina

La función de densidad de probabilidad gaussiana (p) de una variable aleatoria (z) viene dada por:

$$p_G(z) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(z-\mu)^2}{2\sigma^2}}$$

Siendo: σ = desviación estándar y μ = valor medio.

Este tipo de ruido es causado por fuentes naturales como la vibración térmica de los átomos y la naturaleza discreta de los objetos cálidos:

En imágenes digitales, las principales fuentes de ruido gaussiano surgen durante la adquisición de la imagen (como por ejemplo ruido en el sensor causado por una insuficiente iluminación, altas temperaturas) o durante la transmisión de la imagen (los circuitos eléctricos conectados al sensor aportan su propia parte del ruido).

Un ejemplo típico de ruido Gaussiano en imágenes es el llamado ruido de Johnson- Nyquist (ruido térmico) el cual es aditivo, independiente en cada píxel e independiente de la intensidad de la señal.

Durante el procesamiento de imágenes digitales, este ruido puede reducirse utilizando filtros espaciales, sin embargo, la mayoría de las técnicas empleadas para reducir o eliminar el ruido terminan suavizando la imagen y afectando sus características (Figura 4.35).

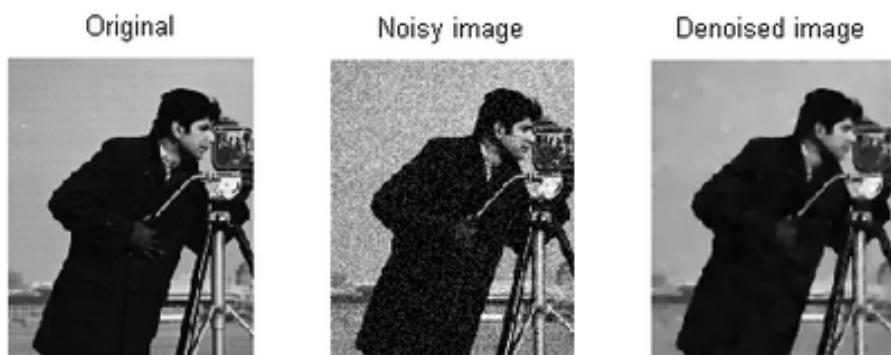


Figura 4.35: De izquierda a derecha: Imagen original, imagen con ruido Gaussiano, imagen sin ruido

4.6.3. Riciano:

Las imágenes de resonancia magnética se ven afectadas por el ruido de Riciano[46], las señales con menos intensidades ($SNR < 2$) están sesgadas debido al ruido. Este ruido subyacente se puede estimar y se puede reducir el factor de sesgo. El ruido de Riciano se forma a partir del ruido gaussiano complejo basado en mediciones en el dominio de frecuencia original.

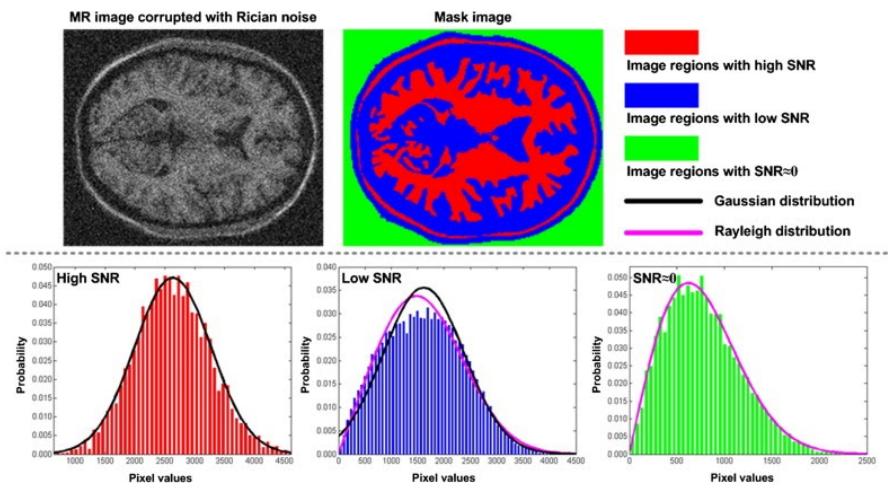


Figura 4.36

Las imágenes de magnitud son más comunes en la RM porque evitan el problema de los artefactos de fase al descartar deliberadamente la información de fase. La señal se mide a través de un detector de cuadratura que proporciona las señales reales e imaginarias. Asumiremos que el ruido en cada señal tiene una distribución gaussiana con media cero y se supondrá que cada canal está contaminado con ruido blanco.

Las imágenes reales e imaginarias se reconstruyen a partir de los datos adquiridos por la compleja transformación de Fourier. Debido a que la transformada de Fourier es una transformación lineal y ortogonal, preservará las características gaussianas del ruido. Además, la varianza del ruido será uniforme en todo el campo de visión y, debido a la transformada de Fourier, el ruido en los véxoles reales e imaginarios correspondientes puede suponerse sin correlación.

Hay muchos factores que influyen en la relación señal-ruido (SNR) final en las imágenes reales e imaginarias. No solo el ruido está asociado con la resistencia de la bobina receptora, sino también con las pérdidas inductivas.

Cuál es la fuente dominante dependerá del campo magnético estático (B_0) y del tamaño del volumen de muestra. Además, el ruido final de la imagen dependerá del tamaño del véxel de la imagen, el ancho de banda del receptor y el número de promedios en la adquisición de la imagen.

La función de densidad de probabilidad es:

$$f(x|\nu, \sigma) = \frac{x}{\sigma^2} \exp\left(\frac{-(x^2 + v^2)}{2\sigma^2}\right) I_0\left(\frac{x\nu}{\sigma^2}\right) \quad (4.1)$$

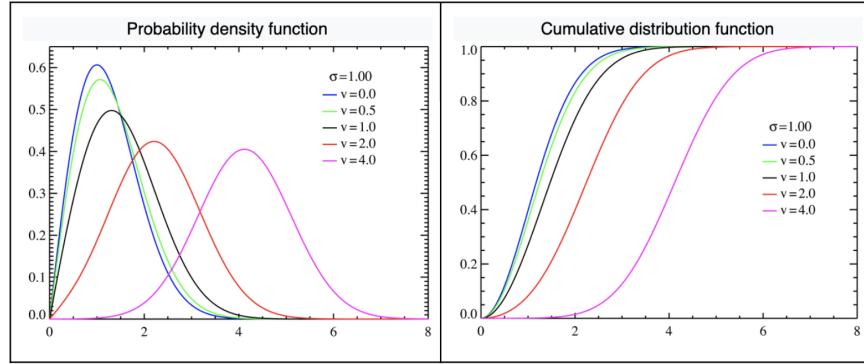


Figura 4.37

5. Metodología

Como se ha visto en el apartado Estado del Arte, actualmente se están utilizando modelos de autoencoders para eliminar ruido en imágenes. A continuación se muestra el tratamiento dado a los datos así como las distintas arquitecturas de modelos que se han usado durante este proyecto.

5.1. Imágenes de Origen

El conjunto de datos fue proporcionado por QUIBIM S.L y está formado por:

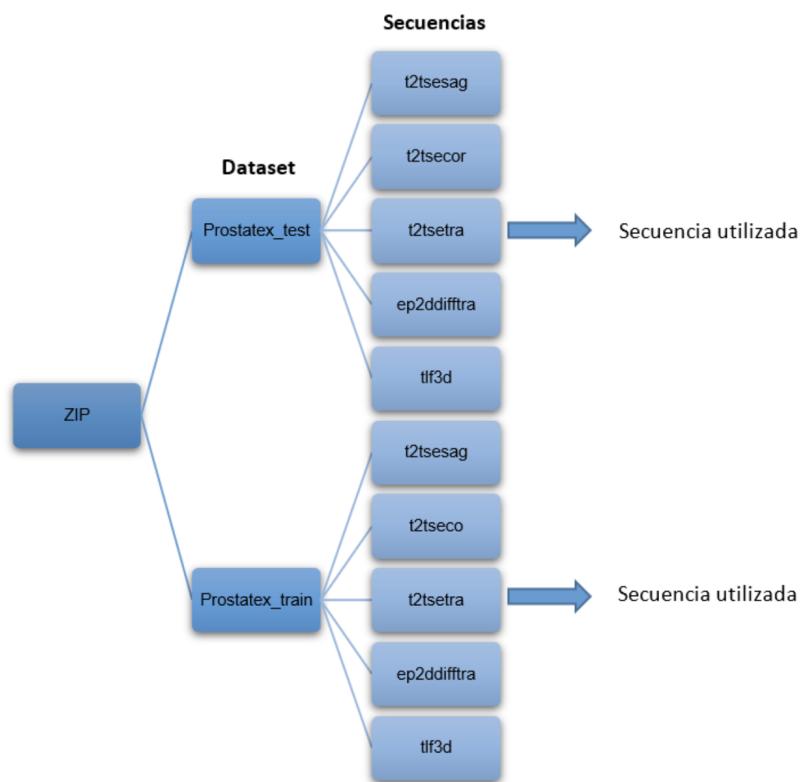


Figura 5.1: Estructura de los datos utilizados

Los datos que contienen los ficheros, son imágenes de resonancia magnética de próstata en formato DICOM. Para la realización de este proyecto, la secuencia de imágenes utilizadas fue **T2 transversal** (técnica en la cual el tiempo de relajación es transversal y depende de la relación entre el protón y los protones vecinos) codificada en el archivo zip como *t2tsetra*.

Las imágenes de resonancia magnética (RM o MRI) que tienen secuencia T2 transversal son imágenes limpias es decir, sin ruido.

La RM en 3D corresponde a la zona pélvica y se puede observar la próstata correspondiente a cada paciente (Figura 5.2).

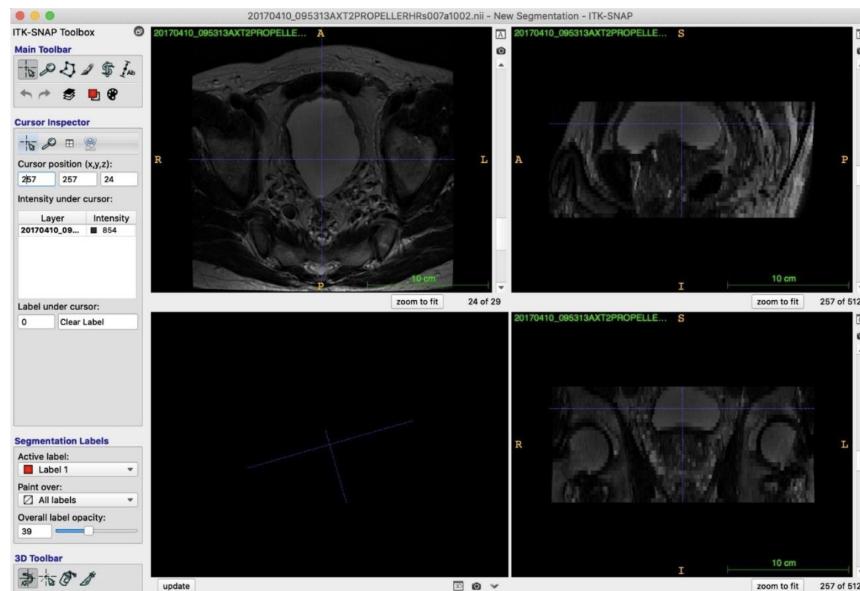


Figura 5.2: Imágenes de próstata secuencia T2 obtenidas con el software ITK-SNAP

5.2. Preprocesado de las Imágenes

Como se ha comentado en el apartado anterior disponemos de imágenes sin ningún tipo de ruido, por tanto, es necesario añadirles ese ruido con la finalidad de entrenar los modelos y conseguir el objetivo deseado (eliminar ruido en imágenes).

5.2.1. Imágenes Originales y Normalización de Píxeles

Las imágenes originales que forman el dataset podemos visualizarlas de la siguiente manera:

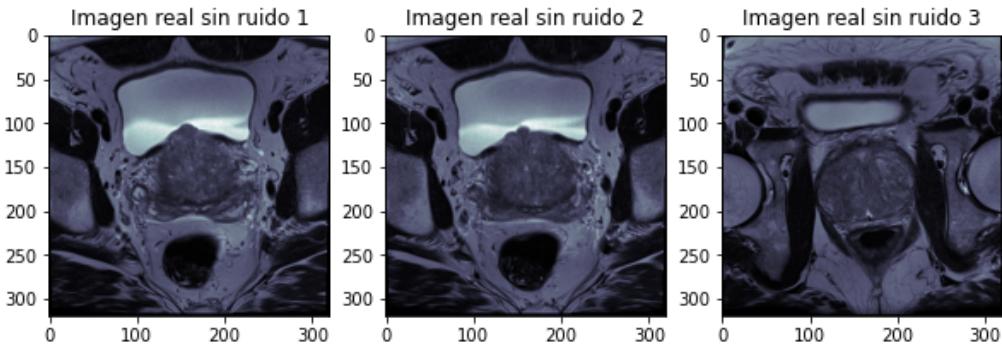


Figura 5.3

Teniendo en cuenta que las imágenes de RM presentan un histograma[47] (Figura 5.4.) en el que se puede identificar dos zonas, la primera de ellas es una distribución de Rayleigh y corresponde al ruido de fondo y la segunda es una gaussiana que se asocia a los píxeles con la información útil, se analizará el histograma obtenido en los datos de partida.

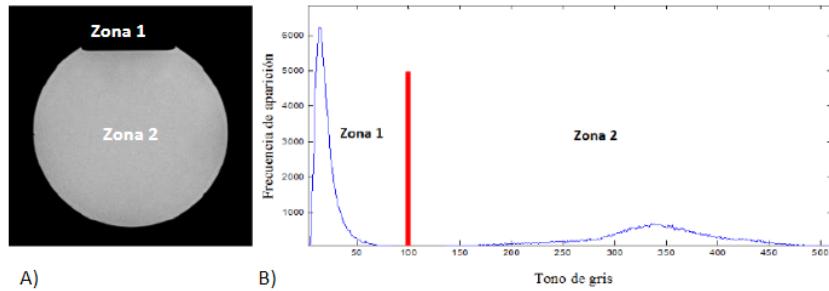


Figura 5.4: a) Imagen de RM de objeto homogéneo. b) Histograma de la imagen .a" que cumple con la distribución de Rice. Este histograma se divide en la distribución de Rayleigh correspondiente al ruido de fondo (Zona 1) y la distribución gaussiana donde se encuentra la información útil (Zona 2)

A continuación se muestra el histograma de los datos de partida sin ningún tipo de normalización (Figura 5.5).

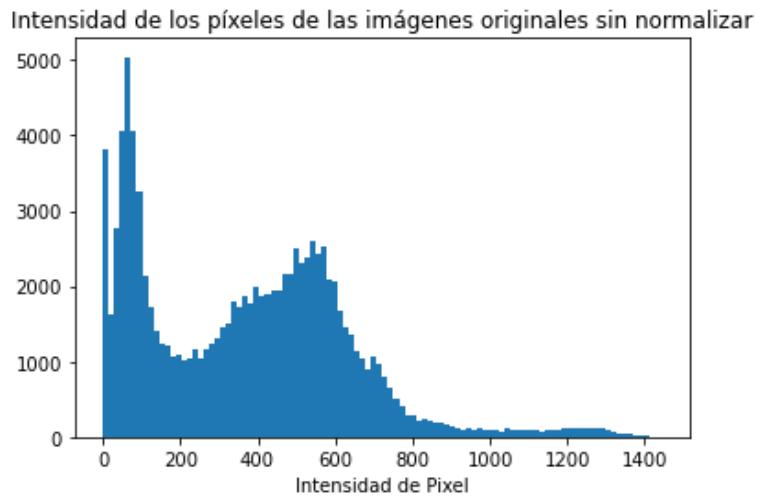


Figura 5.5: Histograma de imágenes originales sin normalizar

En esta imagen se puede observar:

- Ruido de fondo que sigue una distribución Rayleigh.
- Los píxeles con los datos útiles se aproximan a una distribución gaussiana.

Con la finalidad de tener niveles de intensidad comparable en todas las imágenes, se ha realizado una normalización de píxeles teniendo en cuenta todos los datos de la imagen según la siguiente fórmula:

$$X \text{ normalizado} = \frac{X - X_{\min}}{X_{\max} - X_{\min}}$$

El histograma obtenido una vez normalizado los datos es el siguiente:

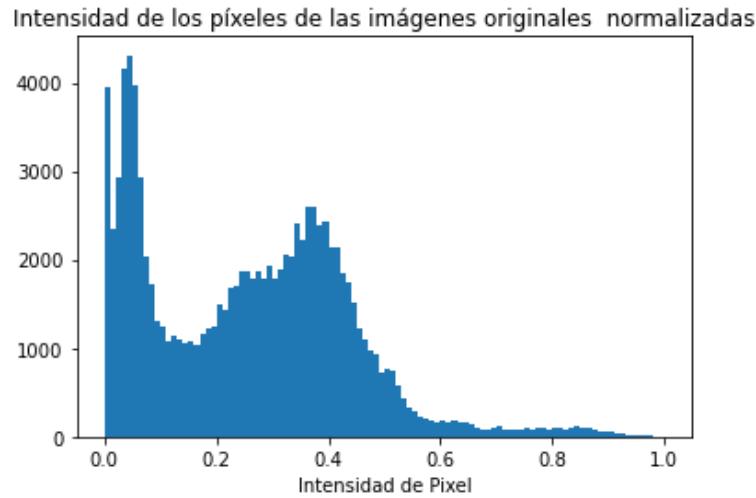


Figura 5.6

Como vemos, no se ha producido alteración en el tipo de distribución de los datos.

5.2.2. Añadir Ruido y Normalización de los Datos

5.2.2.1. Ruido Gaussiano

Durante este proyecto las imágenes originales se han perturbado con dos ruidos Gaussiano, cada uno de ellos presentan las siguientes varianzas ($\text{varianza} = \text{desviacionestndar} * 2$):

- Primer ruido Gauss: varianza = 0.002

Las imágenes obtenidas después de aplicar este ruido, podemos visualizarlas de la siguiente manera:

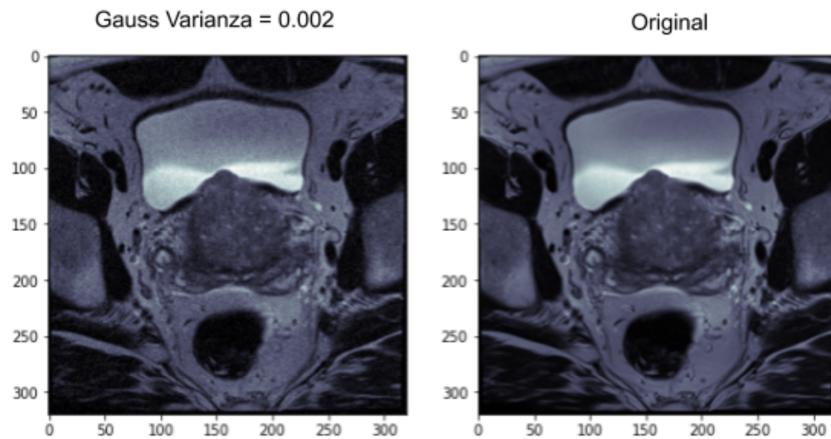


Figura 5.7

El histograma obtenido una vez normalizados los datos, tal y como se ha realizado en las imágenes originales, es el siguiente:

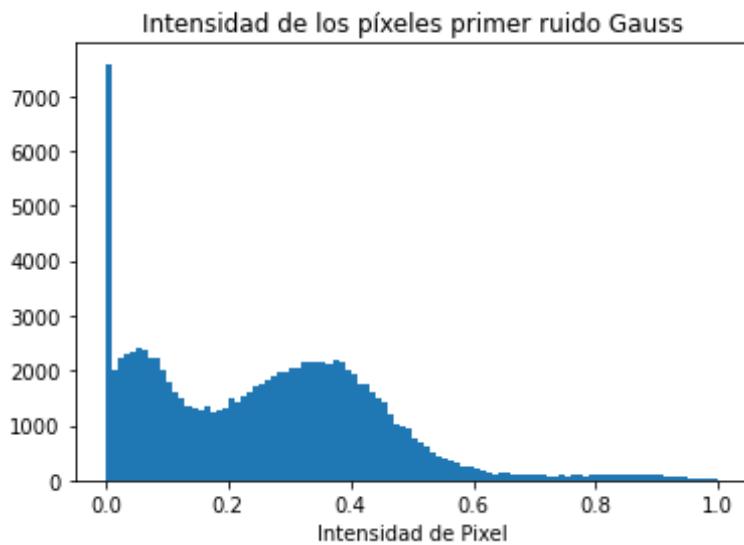


Figura 5.8

En el histograma se observa una distribución Bimodal (tenemos dos grupos de datos distintos, el primero se corresponde con el ruido de fondo y el segundo con la distribución gaussiana donde se encuentra la información útil).

- Segundo ruido Gauss: varianza = 0.01

Las imágenes obtenidas después de aplicar este ruido, podemos visualizarlas de la siguiente manera:

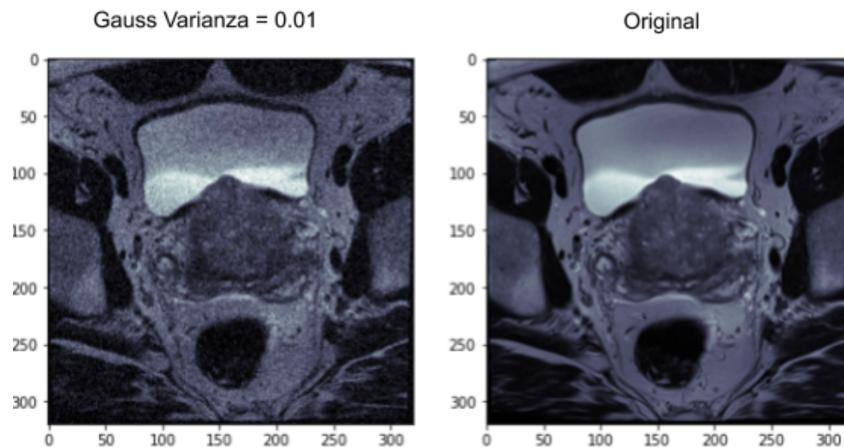


Figura 5.9

Como se puede apreciar en la imagen, el segundo ruido produce una mayor distorsión.

El histograma obtenido una vez normalizados los datos, tal y como se ha realizado en las imágenes originales, es el siguiente:

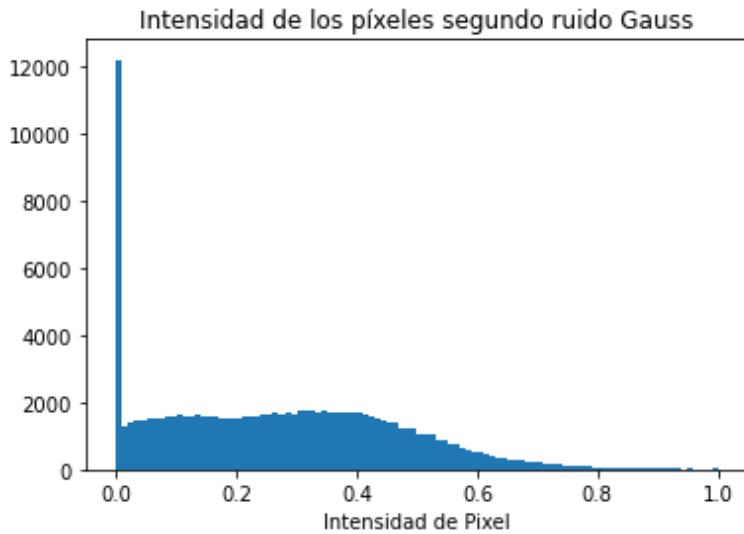


Figura 5.10

En este caso obtenemos una distribución que podríamos aproximar a gaussiana (no se aprecia debido al elevado número de píxeles con intensidad 0).

Además podemos apreciar que conforme aumenta la varianza del ruido, se obtienen histogramas con mayor frecuencia de píxeles cuya intensidad es igual a cero.

5.2.2.2. Ruido Salt and Pepper

En este caso, las imágenes se han perturbado con dos ruidos Salt and Pepper, cada uno de ellos presentan una distinta proporción de píxeles reemplazados con ruido (amount):

- Primer ruido Salt and Pepper: amount= 0.08

Una vez aplicado este ruido, se obtienen las siguientes imágenes:

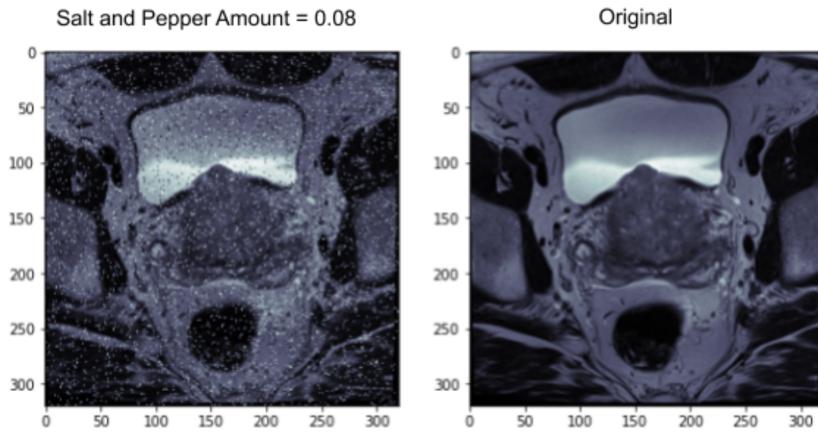


Figura 5.11

A continuación se muestra el histograma obtenido después de aplicar ruido y normalizar las imágenes:

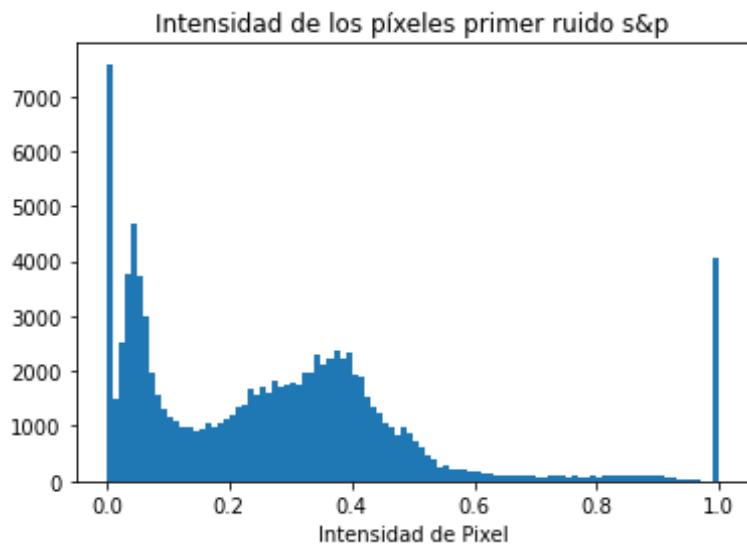


Figura 5.12

Teniendo en cuenta el comportamiento del ruido Salt and Pepper, se han obtenido unos resultados lógicos. La aplicación del ruido ha producido un aumento en la frecuencia de píxeles con intensidad 0 y 1 (puntos negros y blancos).

- Segundo ruido Salt and Pepper: amount= 0.3

Una vez aplicado este ruido, se obtienen las siguientes imágenes:

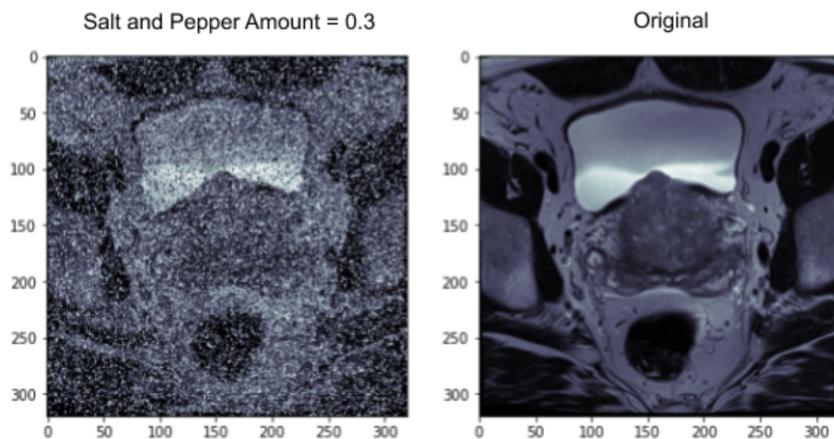


Figura 5.13

Las imágenes obtenidas son más ruidosas que las anteriores y presentan el siguiente histograma:

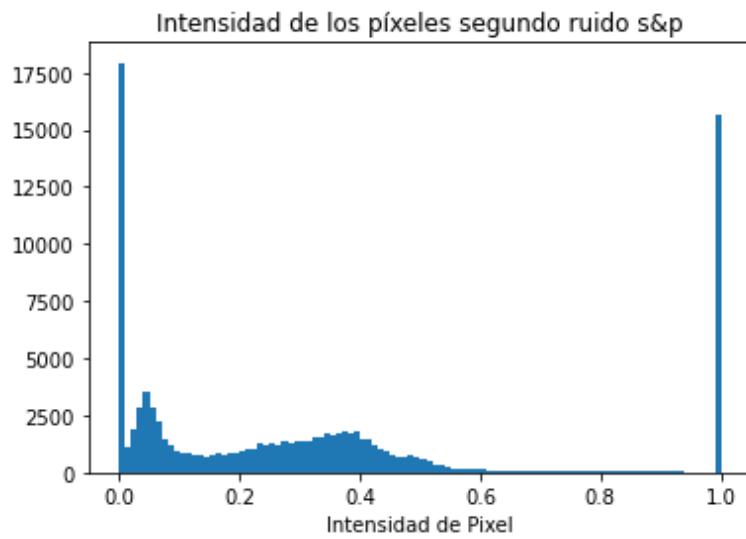


Figura 5.14

Comparando los dos histogramas obtenidos en este ruido se puede apreciar que al aumentar la proporción de píxeles reemplazados con ruido (amount), se produce un aumento en la frecuencia de píxeles cuya intensidad es cero o uno (píxeles negros y blancos respectivamente).

También se puede concluir que aplicar ruido Salt and Pepper en las imágenes no produce cambios en la distribución de píxeles originales.

5.2.2.3. Ruido Riciano

Las imágenes originales se han perturbado con dos ruidos de tipo riciano y cada uno de ellos presenta distintas características:

- Primer ruido Riciano: se ha utilizado una señal-ruido = 18 y señal de referencia = 1.5, obteniéndose el siguiente histograma (los datos se han normalizado).

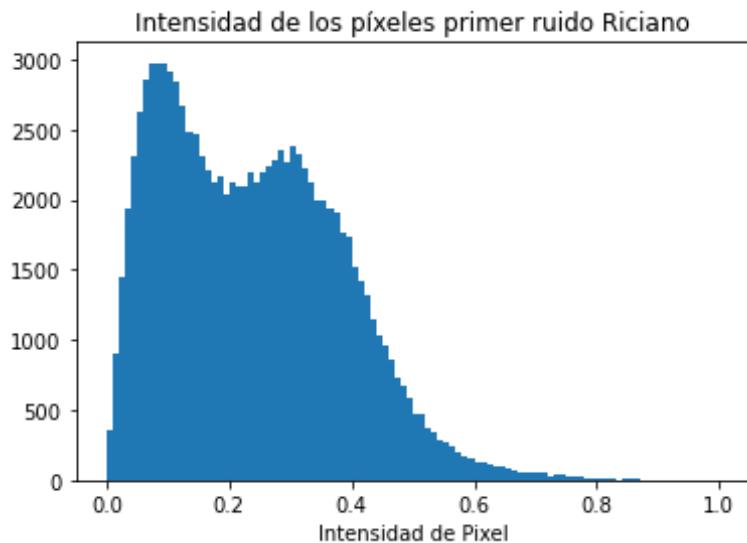


Figura 5.15

Las imágenes obtenidas presentan una distribución binomial. Podemos encontrar dos grupos de píxeles (el primero corresponde al ruido de fondo y el segundo corresponde a la información útil) donde cada uno de ellos se aproxima a una distribución normal.

Las imágenes obtenidas son de la siguiente forma:

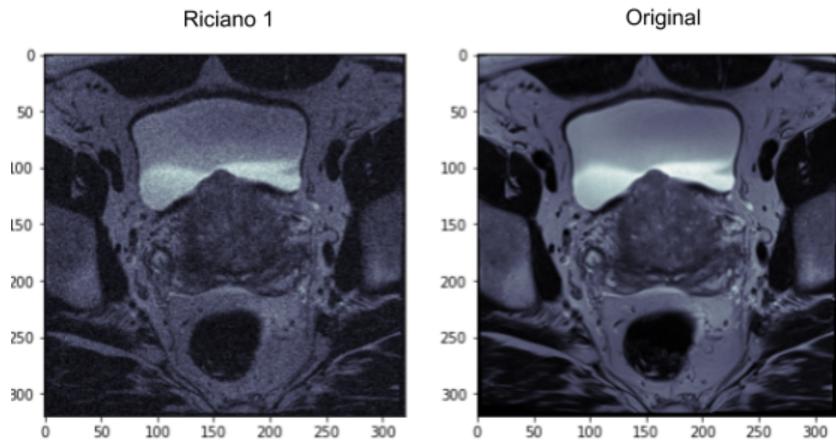


Figura 5.16

- Segundo ruido Ricano: se ha utilizado una señal-ruido = 14 y señal de referencia = 2, obteniéndose el siguiente histograma (los datos se han normalizado).

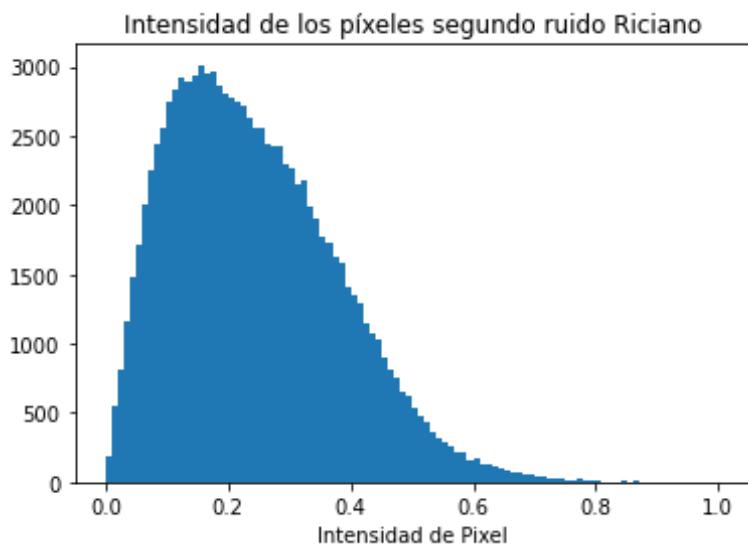


Figura 5.17

En este caso, los datos se distribuyen dando lugar a una distribución gaussiana aproximada.

Las imágenes presentan una mayor perturbación que las obtenidas durante el primer ruido riciano y se visualizan de la siguiente forma:

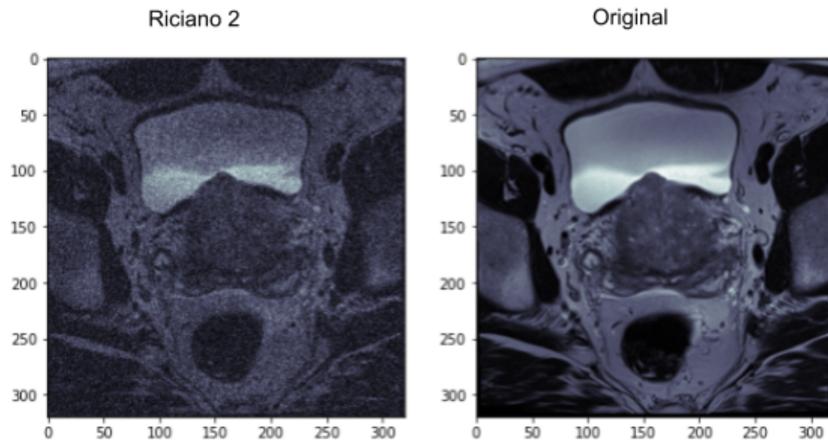


Figura 5.18

5.3. Creación del Conjunto de Datos para el Entrenamiento de los Modelos y su Validación

Una vez realizado el tratamiento de las imágenes, es necesario preparar un conjunto de datos para entrenar los modelos y realizar su correspondiente validación. A continuación se muestra un esquema detallando la creación de los datasets (Figura.5.19).

- Entrenamiento del modelo:

Se parte de un dataset original llamado Prostatex_train con una secuencia T2 transversal. El dataset se divide en 80 % para realizar el entrenamiento y 20 % para testear dicho entrenamiento.

- Validación del modelo:

Se utilizará otro dataset original previamente perturbado con los distintos tipos de ruido llamado Prostatex_test (no ha sido visto por el modelo), para comprobar que los modelos no sobre ajustan, se adaptan a nuevos datos y originan resultados favorables.

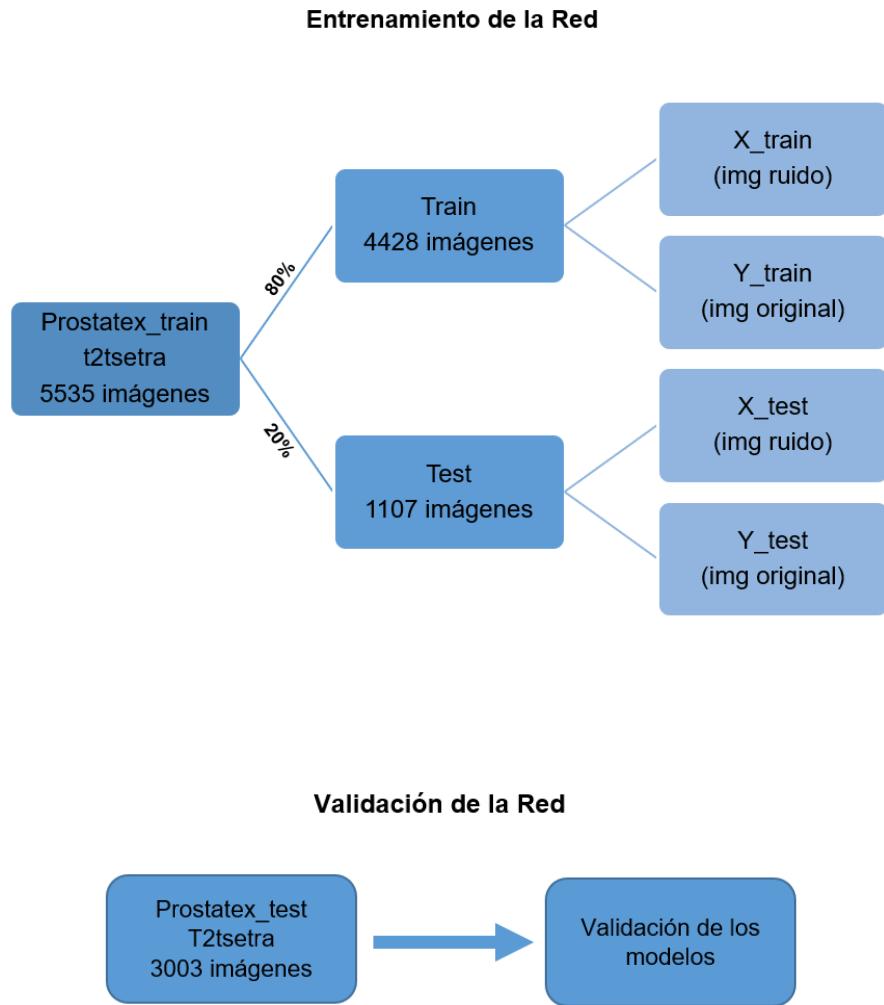


Figura 5.19

5.4. Construcción de los Modelos

Durante la realización de este proyecto, se han utilizado 5 modelos de autoencoders distintos. En las siguientes tablas se pueden observar las diferentes redes:

MODELO 1

Layer (type)	Output Shape	Param #
<hr/>		
input_1 (InputLayer)	(None, 320, 320, 1)	0
conv2d_1 (Conv2D)	(None, 320, 320, 32)	320
max_pooling2d_1 (MaxPooling2)	(None, 160, 160, 32)	0
conv2d_2 (Conv2D)	(None, 160, 160, 64)	18496
max_pooling2d_2 (MaxPooling2)	(None, 80, 80, 64)	0
conv2d_3 (Conv2D)	(None, 80, 80, 128)	73856
conv2d_4 (Conv2D)	(None, 80, 80, 128)	147584
up_sampling2d_1 (UpSampling2)	(None, 160, 160, 128)	0
conv2d_5 (Conv2D)	(None, 160, 160, 64)	73792
up_sampling2d_2 (UpSampling2)	(None, 320, 320, 64)	0
conv2d_6 (Conv2D)	(None, 320, 320, 1)	577
<hr/>		
Total params:	314,625	
Trainable params:	314,625	
Non-trainable params:	0	

MODELO 2

Layer (type)	Output Shape	Param #
<hr/>		
input_1 (InputLayer)	(None, 320, 320, 1)	0
conv2d_1 (Conv2D)	(None, 320, 320, 16)	160
dropout_1 (Dropout)	(None, 320, 320, 16)	0
max_pooling2d_1 (MaxPooling2)	(None, 160, 160, 16)	0
conv2d_2 (Conv2D)	(None, 160, 160, 32)	4640
max_pooling2d_2 (MaxPooling2)	(None, 80, 80, 32)	0
conv2d_3 (Conv2D)	(None, 80, 80, 64)	18496
conv2d_transpose_1 (Conv2DTr)	(None, 160, 160, 64)	36928
conv2d_transpose_2 (Conv2DTr)	(None, 320, 320, 32)	18464
conv2d_transpose_4 (Conv2DTr)	(None, 320, 320, 1)	289
<hr/>		
Total params:	78,977	
Trainable params:	78,977	
Non-trainable params:	0	

MODELO 3

Layer (type)	Output Shape	Param #
<hr/>		
input_2 (InputLayer)	(None, 320, 320, 1)	0
Conv1 (Conv2D)	(None, 320, 320, 32)	320
pool1 (MaxPooling2D)	(None, 160, 160, 32)	0
Conv2 (Conv2D)	(None, 160, 160, 64)	18496
pool2 (MaxPooling2D)	(None, 80, 80, 64)	0
Conv3 (Conv2D)	(None, 80, 80, 64)	36928
upsample1 (UpSampling2D)	(None, 160, 160, 64)	0
Conv4 (Conv2D)	(None, 160, 160, 32)	18464
upsample2 (UpSampling2D)	(None, 320, 320, 32)	0
Conv5 (Conv2D)	(None, 320, 320, 1)	289
<hr/>		
Total params:	74,497	
Trainable params:	74,497	
Non-trainable params:	0	

MODELO 4

Layer (type)	Output Shape	Param #
<hr/>		
input_1 (InputLayer)	(None, 320, 320, 1)	0
Conv1 (Conv2D)	(None, 320, 320, 32)	320
batch_normalization_1 (Batch Normalization)	(None, 320, 320, 32)	128
pool1 (MaxPooling2D)	(None, 160, 160, 32)	0
Conv2 (Conv2D)	(None, 160, 160, 64)	18496
batch_normalization_2 (Batch Normalization)	(None, 160, 160, 64)	256
pool2 (MaxPooling2D)	(None, 80, 80, 64)	0
Conv3 (Conv2D)	(None, 80, 80, 64)	36928
upsample1 (UpSampling2D)	(None, 160, 160, 64)	0
Conv4 (Conv2D)	(None, 160, 160, 32)	18464
upsample2 (UpSampling2D)	(None, 320, 320, 32)	0
Conv5 (Conv2D)	(None, 320, 320, 1)	289
<hr/>		
Total params:	74,881	
Trainable params:	74,689	
Non-trainable params:	192	

MODELO 5: U-Net			
Layer (type)	Output Shape	Param #	Connected to
input_6 (InputLayer)	(None, 320, 320, 1)	0	
batch_normalization_4 (BatchNor	(None, 320, 320, 1)	4	input_6[0][0]
dropout_3 (Dropout)	(None, 320, 320, 1)	0	batch_normalization_4[0][0]
conv2d_14 (Conv2D)	(None, 320, 320, 8)	80	dropout_3[0][0]
conv2d_15 (Conv2D)	(None, 320, 320, 8)	584	conv2d_14[0][0]
max_pooling2d_7 (MaxPooling2D)	(None, 160, 160, 8)	0	conv2d_15[0][0]
conv2d_16 (Conv2D)	(None, 160, 160, 16)	1168	max_pooling2d_7[0][0]
conv2d_17 (Conv2D)	(None, 160, 160, 16)	2320	conv2d_16[0][0]
max_pooling2d_8 (MaxPooling2D)	(None, 80, 80, 16)	0	conv2d_17[0][0]
conv2d_18 (Conv2D)	(None, 80, 80, 32)	4640	max_pooling2d_8[0][0]
conv2d_19 (Conv2D)	(None, 80, 80, 32)	9248	conv2d_18[0][0]
max_pooling2d_9 (MaxPooling2D)	(None, 40, 40, 32)	0	conv2d_19[0][0]
conv2d_20 (Conv2D)	(None, 40, 40, 64)	18496	max_pooling2d_9[0][0]
conv2d_21 (Conv2D)	(None, 40, 40, 64)	36928	conv2d_20[0][0]
max_pooling2d_10 (MaxPooling2D)	(None, 20, 20, 64)	0	conv2d_21[0][0]
conv2d_22 (Conv2D)	(None, 20, 20, 128)	73856	max_pooling2d_10[0][0]
conv2d_23 (Conv2D)	(None, 20, 20, 128)	147584	conv2d_22[0][0]
conv2d_transpose_6 (Conv2DTrans	(None, 40, 40, 64)	32832	conv2d_23[0][0]
concatenate_1 (Concatenate)	(None, 40, 40, 128)	0	conv2d_transpose_6[0][0]
conv2d_24 (Conv2D)	(None, 40, 40, 64)	73792	concatenate_1[0][0]
conv2d_25 (Conv2D)	(None, 40, 40, 64)	36928	conv2d_24[0][0]
conv2d_transpose_7 (Conv2DTrans	(None, 80, 80, 32)	8224	conv2d_25[0][0]
concatenate_2 (Concatenate)	(None, 80, 80, 64)	0	conv2d_transpose_7[0][0]
conv2d_26 (Conv2D)	(None, 80, 80, 32)	18464	concatenate_2[0][0]
conv2d_27 (Conv2D)	(None, 80, 80, 32)	9248	conv2d_26[0][0]
conv2d_transpose_8 (Conv2DTrans	(None, 160, 160, 16)	2064	conv2d_27[0][0]
concatenate_3 (Concatenate)	(None, 160, 160, 32)	0	conv2d_transpose_8[0][0]
conv2d_28 (Conv2D)	(None, 160, 160, 16)	4624	concatenate_3[0][0]
conv2d_29 (Conv2D)	(None, 160, 160, 16)	2320	conv2d_28[0][0]
conv2d_transpose_9 (Conv2DTrans	(None, 320, 320, 8)	520	conv2d_29[0][0]
concatenate_4 (Concatenate)	(None, 320, 320, 16)	0	conv2d_transpose_9[0][0]
conv2d_30 (Conv2D)	(None, 320, 320, 8)	1168	concatenate_4[0][0]
conv2d_31 (Conv2D)	(None, 320, 320, 8)	584	conv2d_30[0][0]
conv2d_32 (Conv2D)	(None, 320, 320, 1)	9	conv2d_31[0][0]
Total params:	485,677		
Trainable params:	485,675		
Non-trainable params:	2		

5.5. Métricas de Validación

Todos los modelos necesitan una medida o coeficiente que sean capaces de medir la calidad de los resultados obtenidos.

Teniendo en cuenta que durante este proyecto como variable de entrada en el entrenamiento de las redes fueron las imágenes perturbadas con los distintos ruidos, mientras que, las variables de salida fueron las imágenes sin perturbación, tenemos modelos de regresión por lo que se utilizó como métrica el coeficiente de determinación (R^2).

5.5.1. Coeficiente de Determinación (R^2)

El **coeficiente de correlación elevado al cuadrado** (r^2) se denomina *coeficiente de determinación*[48]. Tiene como objetivo medir cómo de bueno es un modelo que se conoce como “*bondad del ajuste*” un ajuste que puede ser lineal o no y que además puede tener distintas variables explicativas. El coeficiente de determinación indica la proporción (o porcentaje, multiplicando por 100) de variabilidad común entre las variables, en otras palabras, indica la proporción de varianza de una variable determinada o asociada a la otra variable. En términos más simples, (r^2) indica el tanto por ciento de acuerdo o de variabilidad común entre ambas variables.

Es importante saber que el resultado del coeficiente de determinación oscila entre 0 y 1. Cuanto más cerca de 1 se sitúe su valor, mayor será el ajuste del modelo a la variable que estamos intentando explicar. De forma inversa, cuanto más cerca de cero, menos ajustado estará el modelo y, por tanto, menos fiable será. La fórmula que expresa esta medida es la siguiente:

$$R^2 = \frac{\sum_{t=1}^T (\hat{Y}_t - \bar{Y})^2}{\sum_{t=1}^T (Y_t - \bar{Y})^2}$$

5.5.2. Coeficiente de Determinación Ajustado

El **coeficiente de determinación ajustado** (*R cuadrado ajustado*) es la medida que define el porcentaje explicado por la varianza de la regresión en relación con la varianza de la variable explicada. Es decir, lo mismo que el R cuadrado, pero con una diferencia. Esa diferencia se encuentra en que el coeficiente de determinación ajustado penaliza la inclusión de variables.

Cómo hemos dicho anteriormente el coeficiente de determinación de un modelo aumenta aunque las variables que incluyamos no sean relevantes. Ya que esto supone un problema, para intentar solventarlo el R cuadrado ajustado queda de la siguiente manera:

$$\bar{R}^2 = \frac{N - 1}{N - k - 1} [1 - R^2]$$

5.6. Herramientas y Tecnología

Software Utilizado:

- **GUI y Gestor de Entorno:**
 - Anaconda Navigator 4.7.10
- **Lenguaje de programación:**
 - Python 3.7
- **Entorno de trabajo/IDEs:**
 - Spyder 3.3.6
 - Google Colaboratory

Librerías:

- **Cálculo Numérico y Análisis de Datos:**
 - Pandas 1.0
 - Numpy 1.18
- **Tratamiento de Imágenes:**
 - Skimage de scikit-image 0.16
 - Dipy 1.1
 - Pydicom 1.4
- **Visualización:**
 - Matplotlib 3.2
- **ML y DL Aplicado en los Modelos:**
 - Scikit-learn 0.22
 - Keras 2.3
- **Específicas para Trabajar en Google Drive:**
 - Googledrivedownloader
 - Google.colab método files

6. Resultados

En este apartado se explica los resultados obtenidos en los entrenamientos de los modelos así como la validación de los mismos.

A continuación se representa de forma gráfica los distintos tipos de ruido utilizados para perturbar las imágenes de entrada a los modelos:

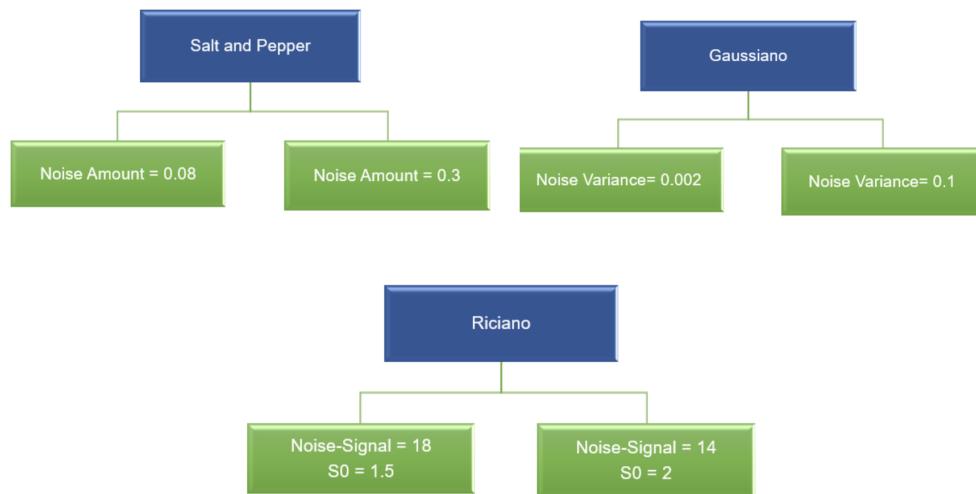


Figura 6.1

6.1. Ruido Salt and Pepper

6.1.1. Primer Ruido Salt and Pepper

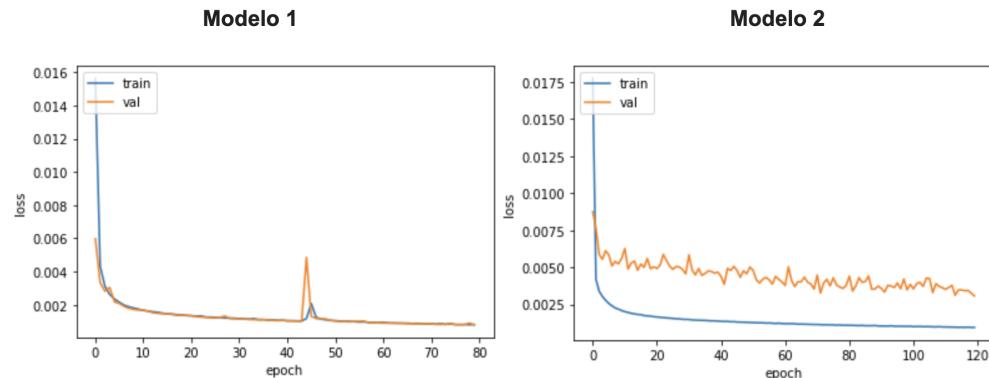
- Primer ruido Salt and Pepper: amount= 0.08

6.1.1.1. Entrenamiento de los Modelos

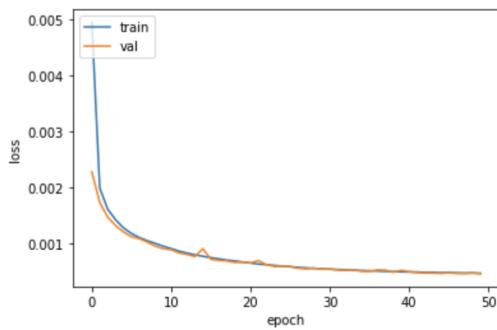
A continuación se adjunta una tabla en la cual se puede encontrar distintos parámetros característicos para cada modelo:

Salt and Pepper: Amount = 0.08				
Modelos	Num.Epcas	Batch_size	T.Entrenamiento (min)	R2 Score
Modelo 1	80	70	39	0.98
Modelo 2	120	70	44	0.85
Modelo 3	50	10	16	0.99
Modelo 4	40	5	18	0.99
Modelo 5: U-Net	100	25	42	-0.24

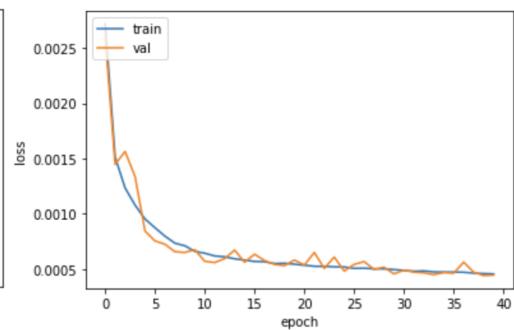
A continuación se muestra las distintas gráficas de entrenamiento (Plot Training and Validation Loss values):



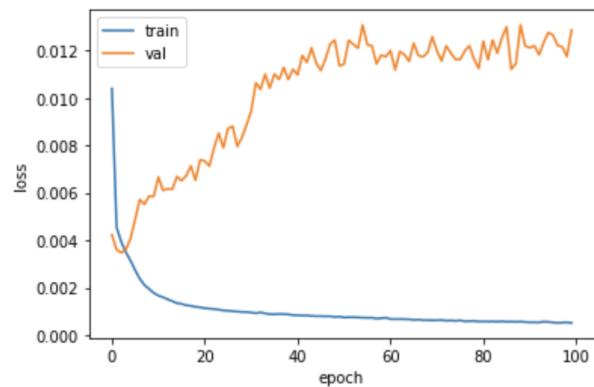
Modelo 3



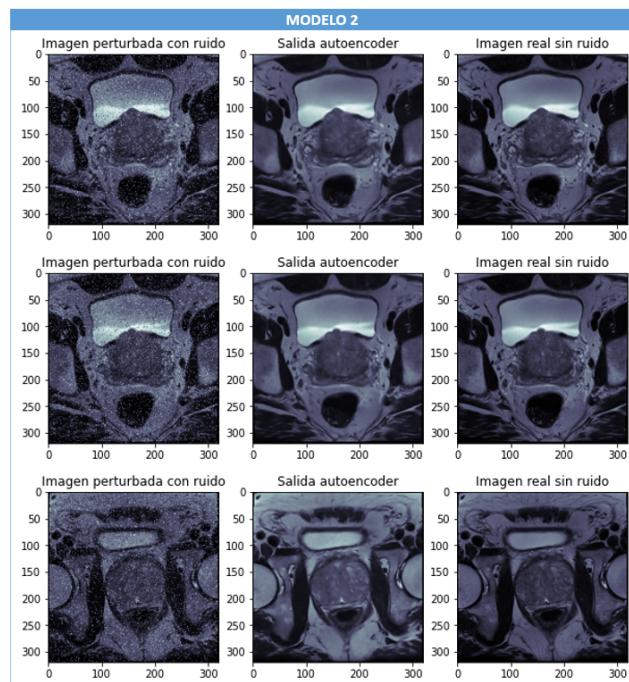
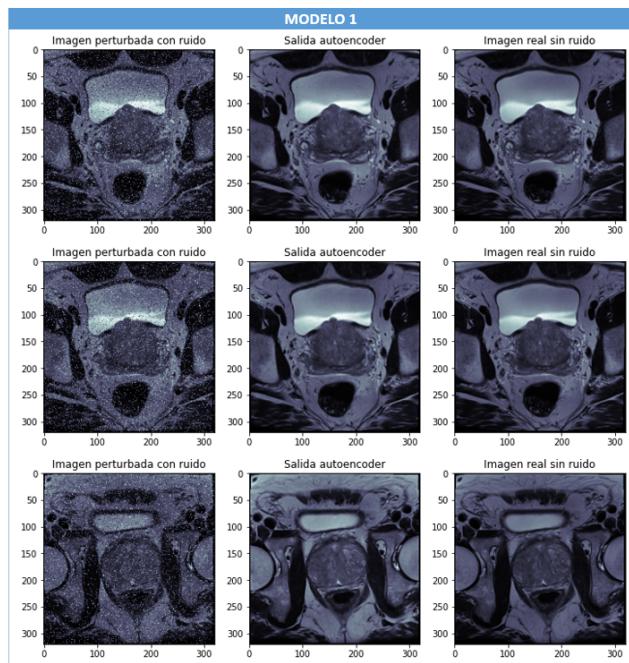
Modelo 4

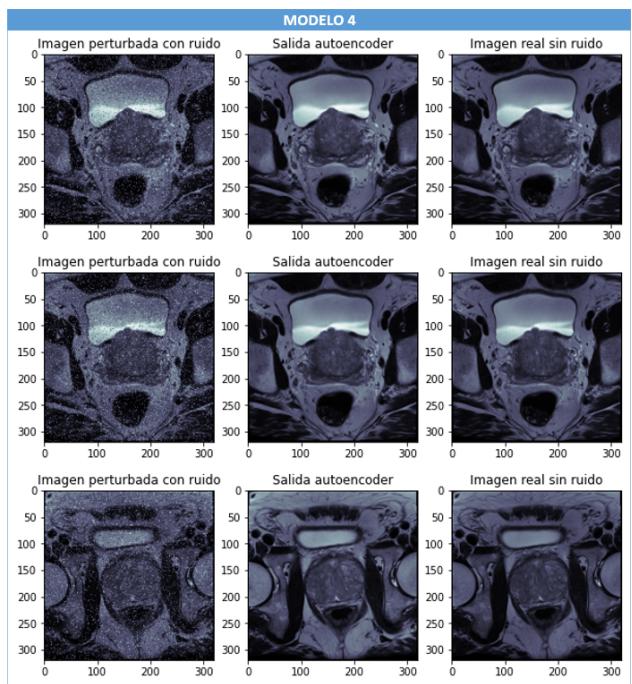
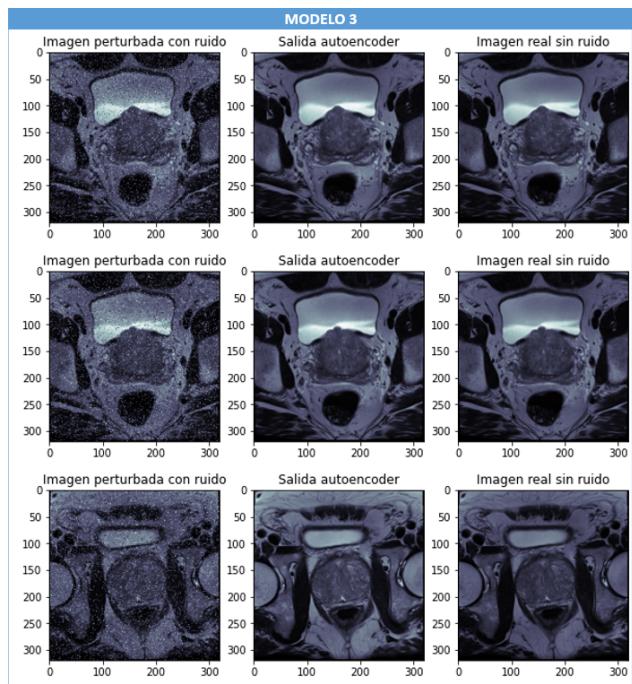


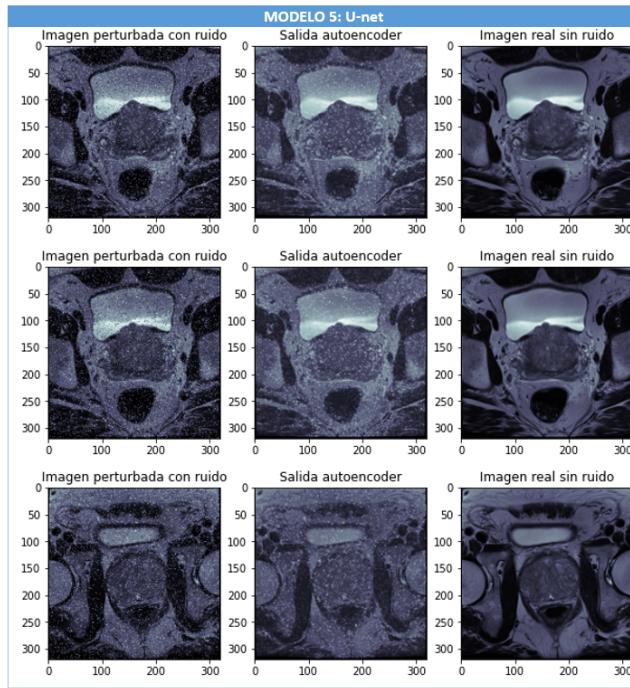
Modelo 5 U-Net



Las imágenes obtenidas durante el entrenamiento de los distintos modelos son:







Observando los resultados obtenidos de las distintas redes podemos concluir que los modelos 1,3 y 4 son los que presentan resultados más favorables.

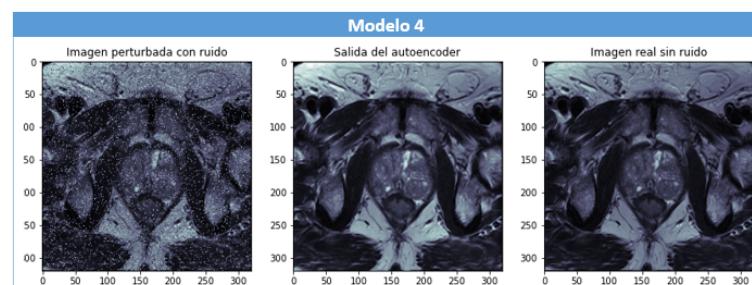
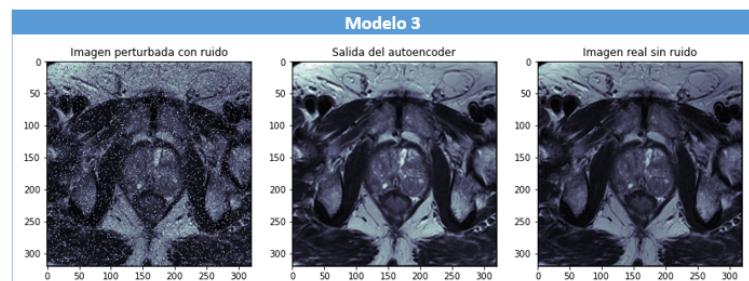
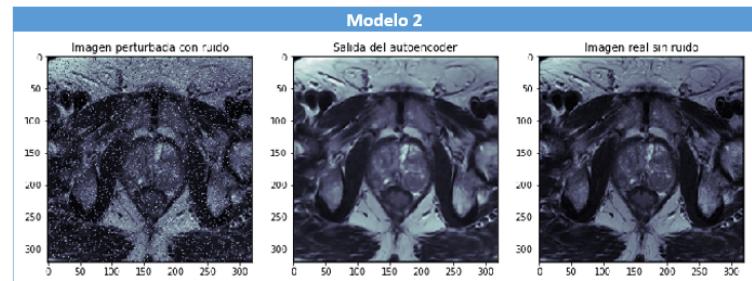
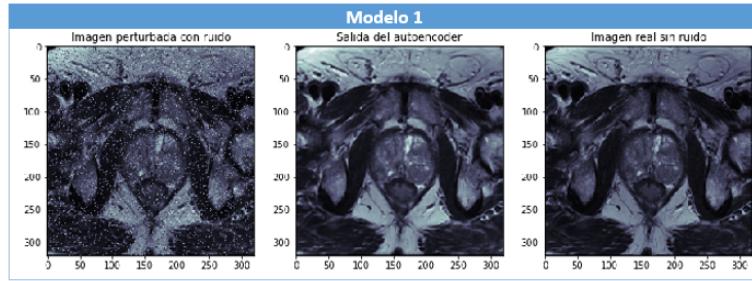
Estas redes han sido capaces de eliminar el ruido de las imágenes de entrada dando como resultado imágenes sin perturbaciones.

Por el contrario, los modelos 2 y 5 (U-Net) no han sido capaces de reconstruir las imágenes manteniendo una similitud a las originales de entrada sin ruido.

6.1.1.2. Validación de los Modelos

En este apartado se han validado los modelos entrenados con nuevos datos no vistos previamente por las redes. A continuación se exponen los resultados obtenidos:

Modelos	R ² Entrenamiento	R ² Validación
Modelo 1	0.98	0.94
Modelo 2	0.85	0.82
Modelo 3	0.99	0.97
Modelo 4	0.99	0.97
U-Net	-0.24	--



Durante la fase de validación de las distintas redes se observa una disminución poco significativa en los valores de R2 y en cuanto a la capacidad de denoising obtenemos resultados

positivos, es decir, las imágenes de salida no presentan ruido.

Sin embargo, se puede observar una pequeña reducción en los detalles presentes en las imágenes iniciales pero se mantiene una definición general positiva.

6.1.2. Segundo Ruido Salt and Pepper

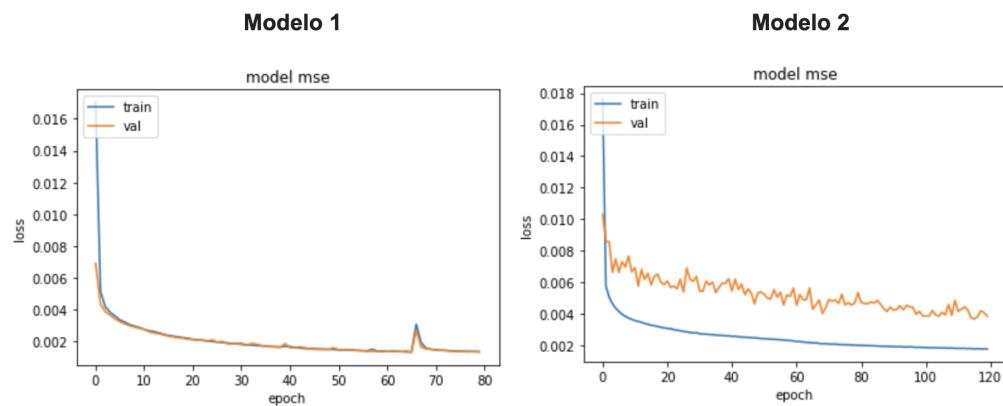
- Segundo ruido Salt and Pepper: amount= 0.3

6.1.2.1. Entrenamiento de los Modelos

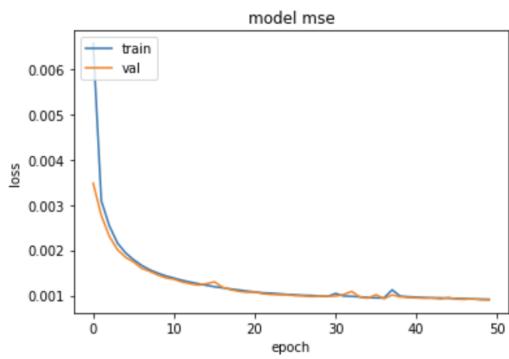
A continuación se adjunta una tabla en la cual se puede encontrar distintos parámetros característicos para cada modelo:

Salt and Pepper: Amount = 0.3				
Modelos	Num.Epocas	Batch_size	T.Entrenamiento (min)	R2 Score
Modelo 1	80	70	39	0.96
Modelo 2	120	70	44	0.79
Modelo 3	50	10	17	0.97
Modelo 4	40	5	18	0.97
Modelo 5: U-Net	100	25	42	0.50

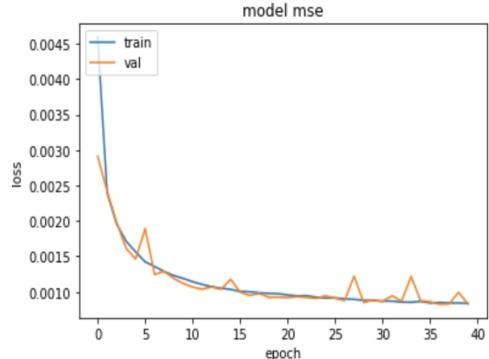
Plot Training and Validation Loss values:



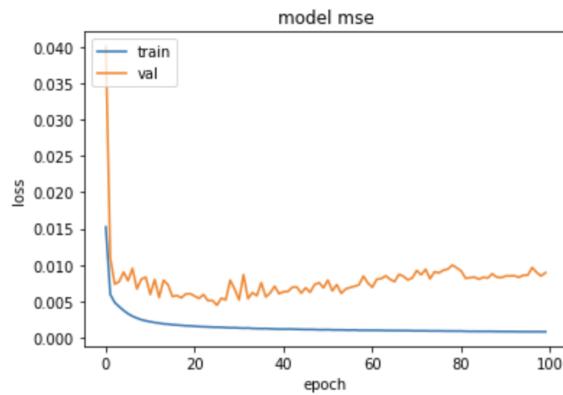
Modelo 3



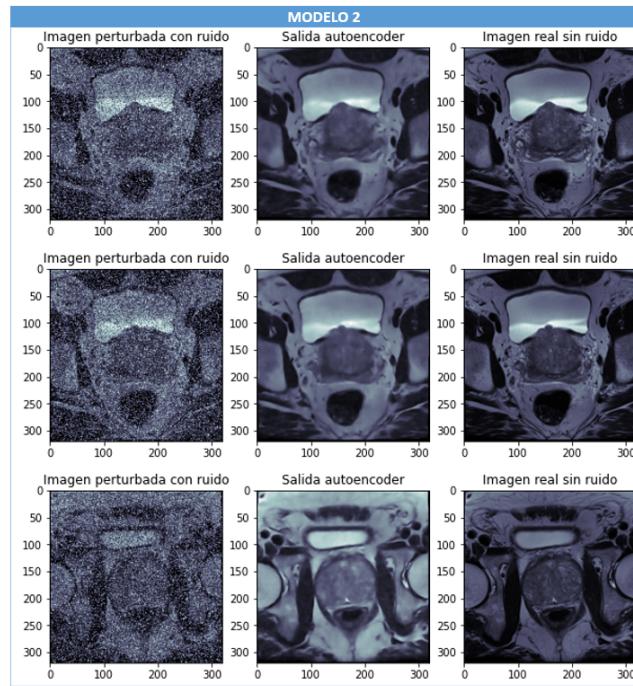
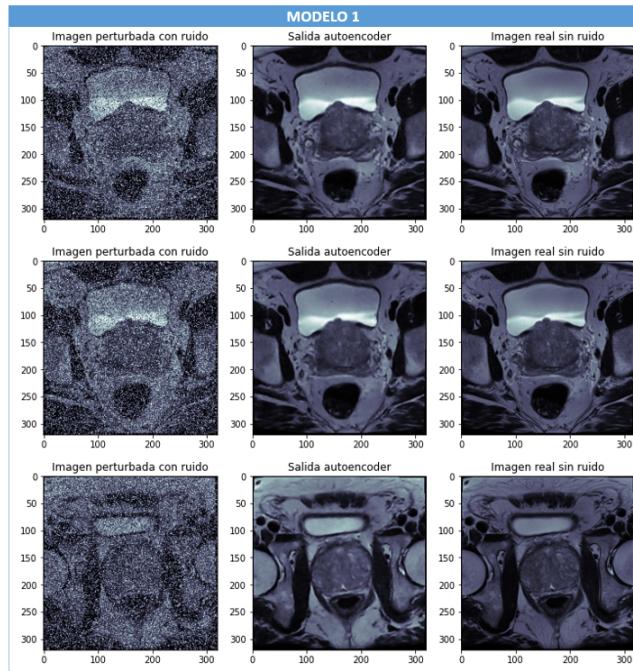
Modelo 4

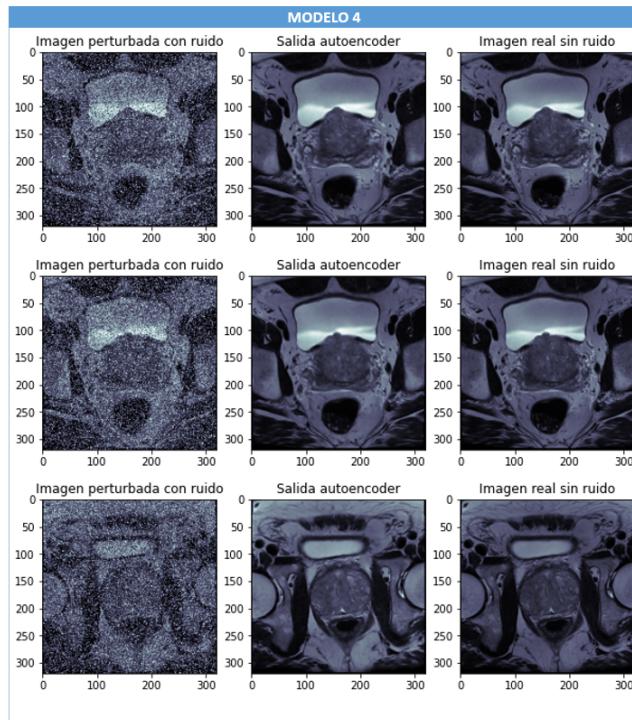
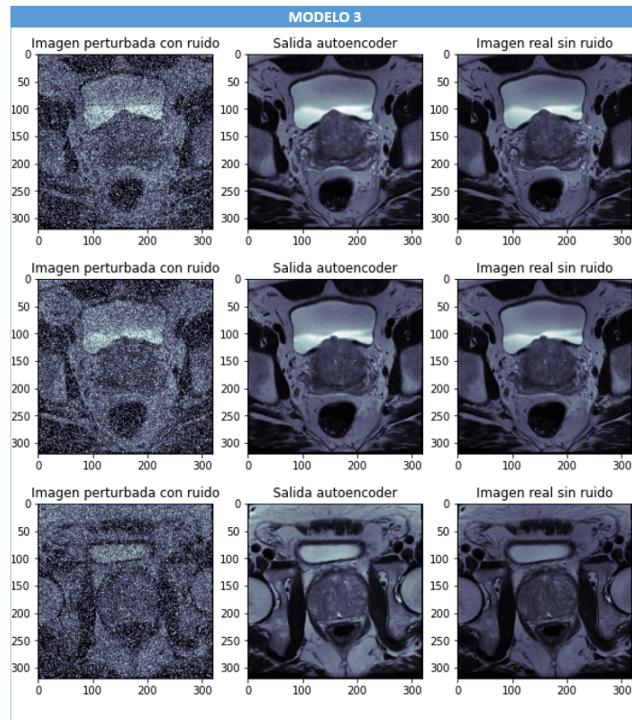


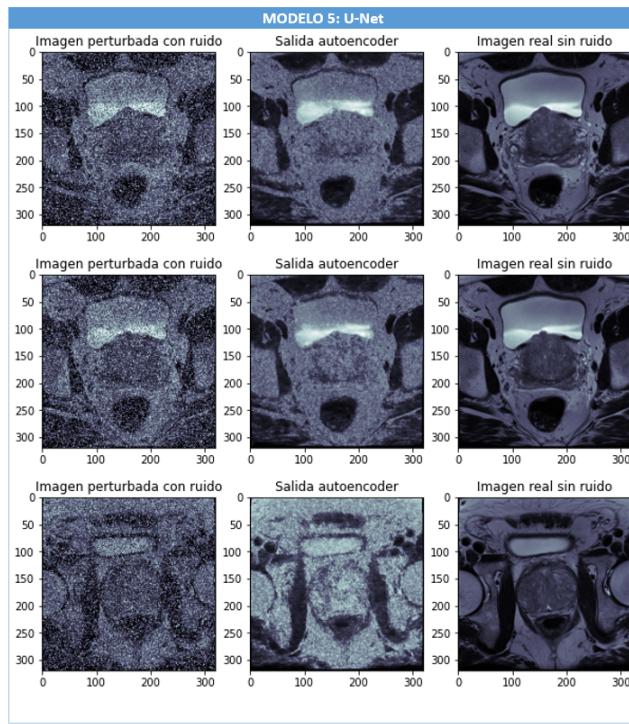
U-Net



Las imágenes obtenidas durante el entrenamiento de los distintos modelos son:







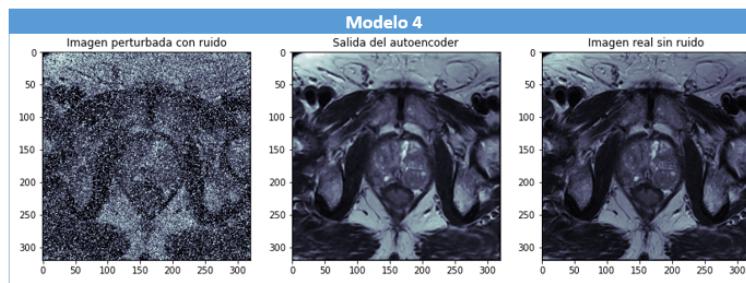
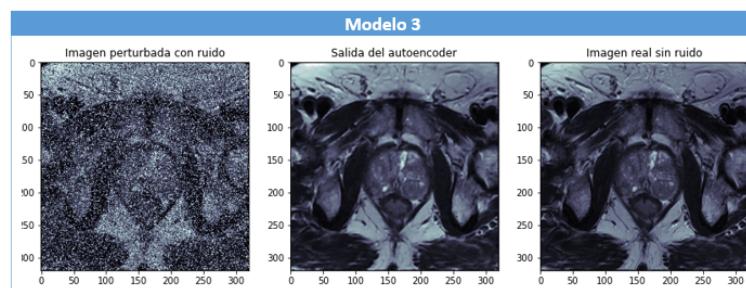
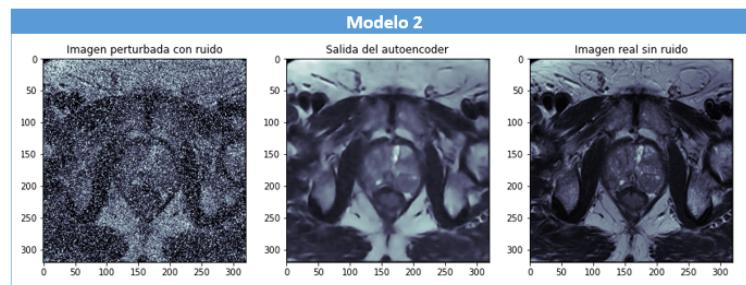
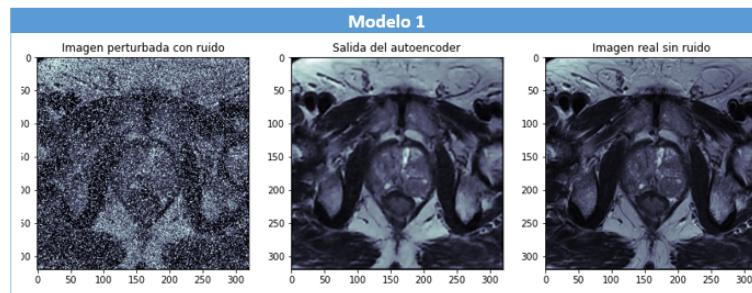
Los resultados obtenidos son similares a la fase de entrenamiento del anterior ruido. Los modelos 1,3 y 4 son los que presentan resultados más favorables.

Por el contrario, los modelos 2 y 5 (U-Net) no han sido capaces de reconstruir las imágenes manteniendo una similitud a las originales de entrada sin ruido.

6.1.2.2. Validación de los Modelos

Los resultados obtenidos durante la validación son:

Modelos	R ² Entrenamiento	R ² Validación
Modelo 1	0.96	0.93
Modelo 2	0.79	0.72
Modelo 3	0.97	0.95
Modelo 4	0.97	0.95
U-Net	0.50	--



Tal y como ocurre en el modelo anterior, no se ha producido una disminución significativa en los valores de R2. Además las redes a pesar del aumento en la densidad de perturbación han sido capaces de eliminar el ruido introducido manteniendo una definición general positiva de las imágenes.

6.2. Ruido Gaussiano

6.2.1. Primer Ruido Gaussiano

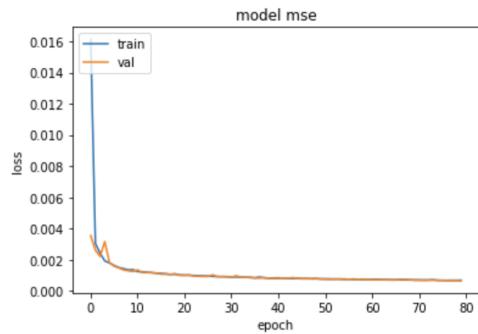
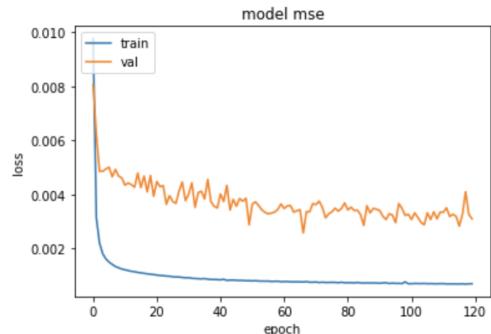
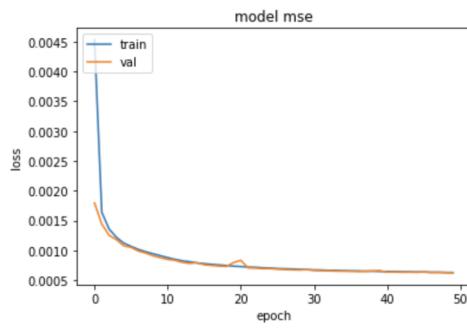
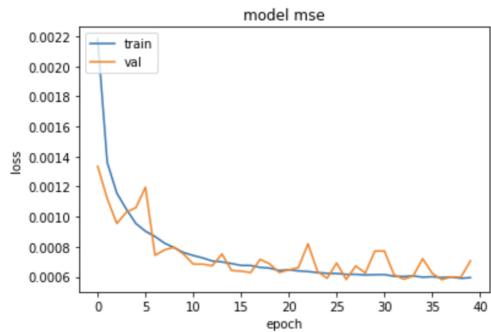
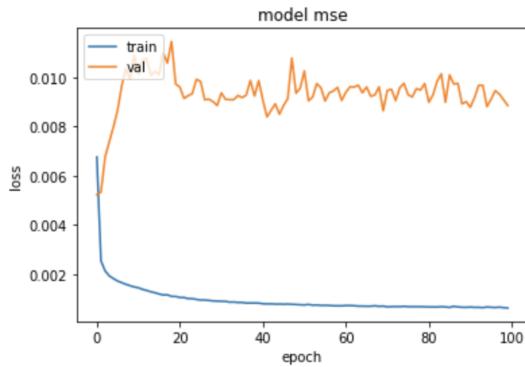
- Primer ruido Gaussiano: variance= 0.002

6.2.1.1. Entrenamiento de los Modelos

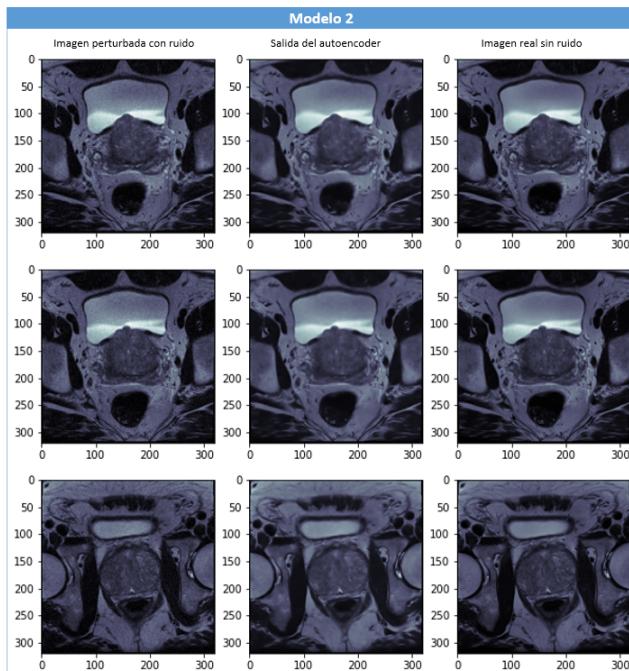
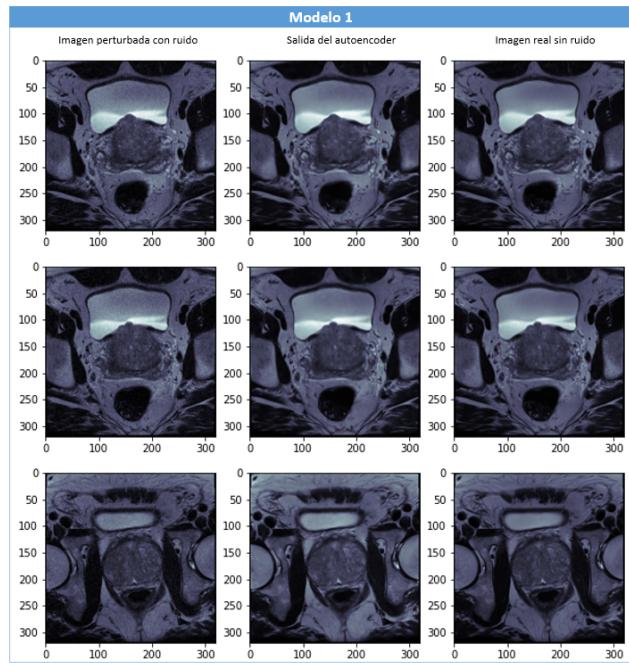
A continuación se adjunta una tabla en la cual se puede encontrar distintos parámetros característicos para cada modelo:

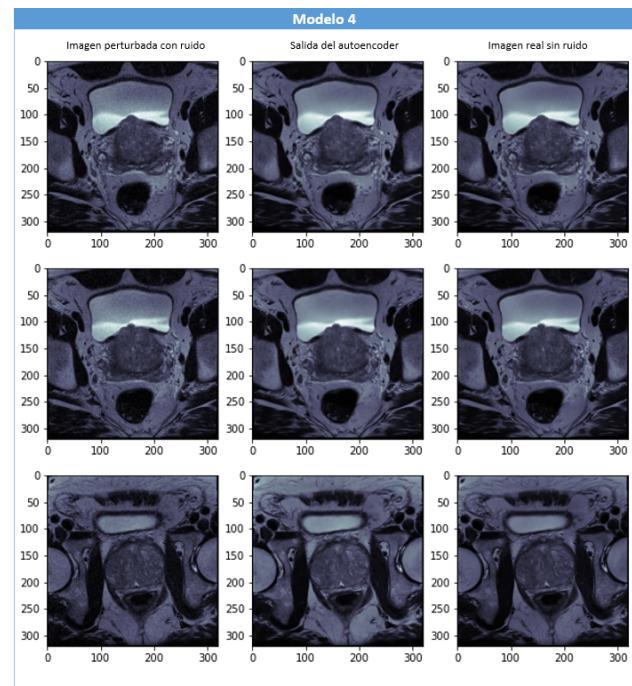
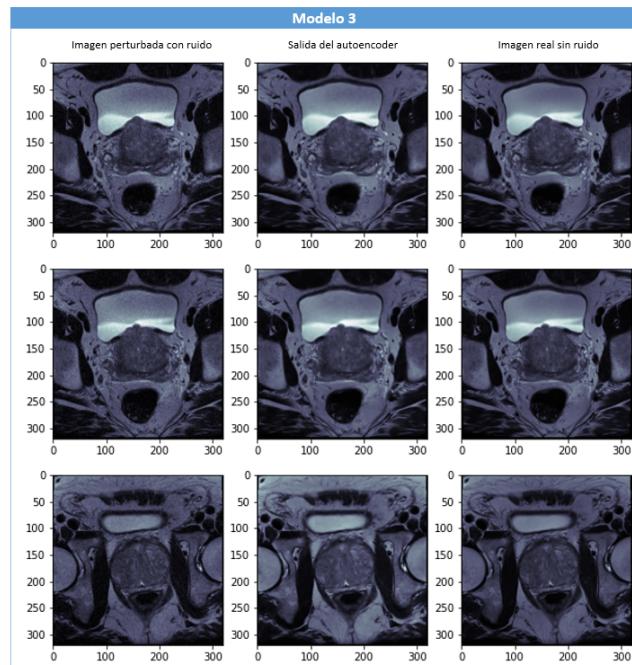
Gaussiano: Variance = 0.002				
Modelos	Num.Epocas	Batch_size	T.Entrenamiento (min)	R2 Score
Modelo 1	80	70	133	0.98
Modelo 2	120	70	130	0.97
Modelo 3	50	10	45	0.98
Modelo 4	40	5	17	0.98
Modelo 5: U-Net	100	25	50	0.05

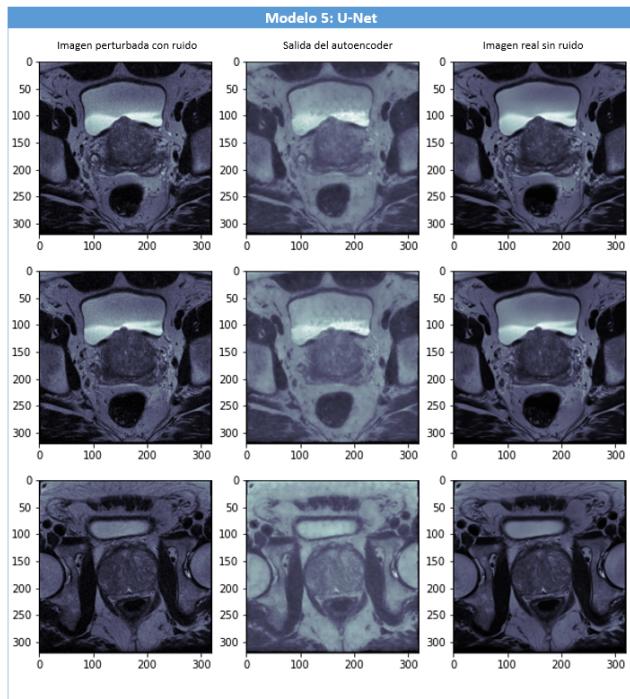
Plot Training and Validation Loss values:

Modelo 1**Modelo 2****Modelo 3****Modelo 4****U-Net**

Las imágenes obtenidas durante el entrenamiento de los distintos modelos son:





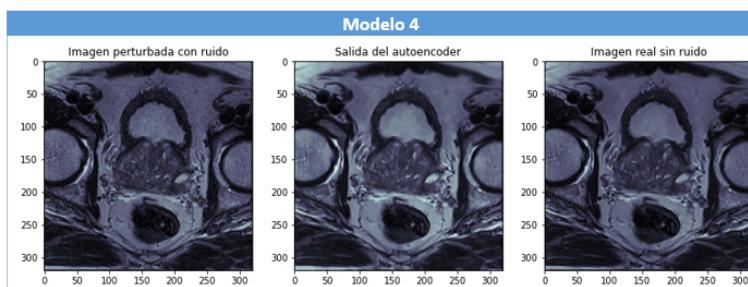
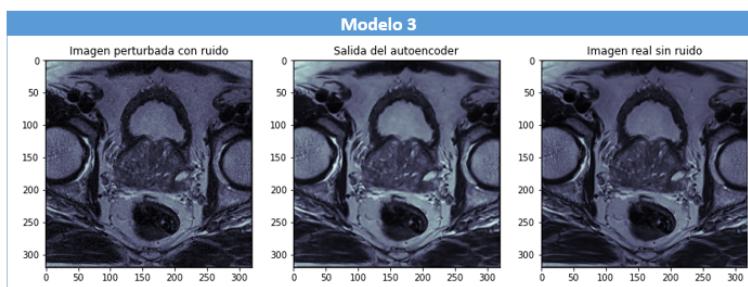
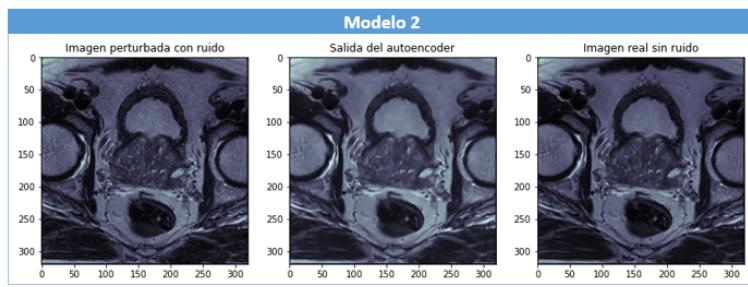
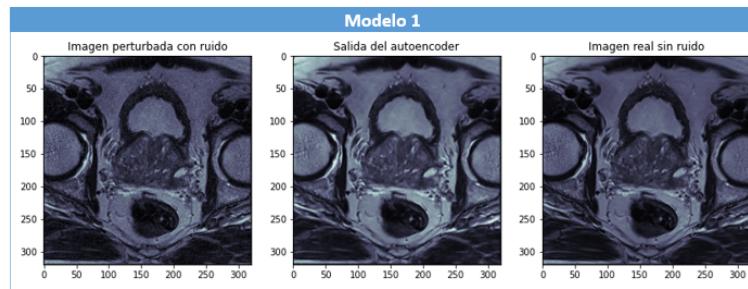


En este caso se observa que todas las redes excepto el modelo U-Net han sido capaces de reducir la perturbación y obtener imágenes de calidad aceptable.

6.2.1.2. Validación de los Modelos

Los resultados obtenidos durante la validación son:

Modelos	R ² Entrenamiento	R ² Validación
Modelo 1	0.98	0.95
Modelo 2	0.97	0.82
Modelo 3	0.98	0.97
Modelo 4	0.98	0.96
U-Net	0.05	--



Los valores de R2 obtenidos en la validación son algo menores que los obtenidos durante el entrenamiento salvo para el modelo 2 donde se ha producido una disminución significativa, lo que nos indica que puede haber sobreajuste durante el entrenamiento del modelo.

6.2.2. Segundo Ruido Gaussiano

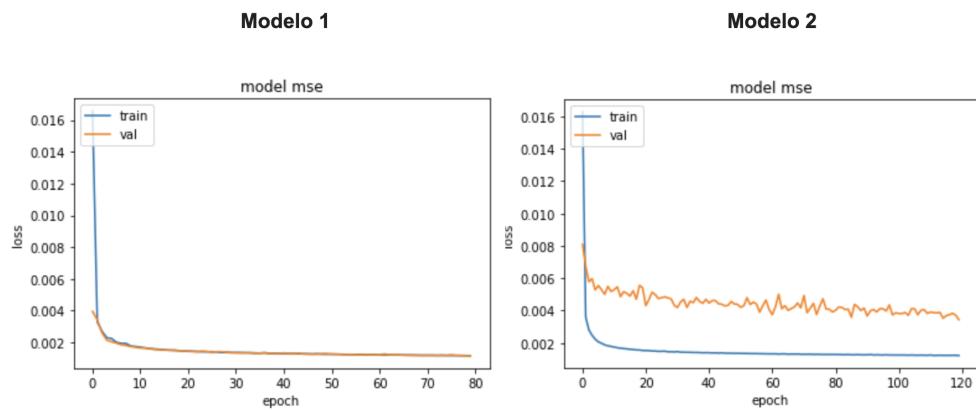
- Segundo ruido Gaussiano: variance= 0.1

6.2.2.1. Entrenamiento de los Modelos

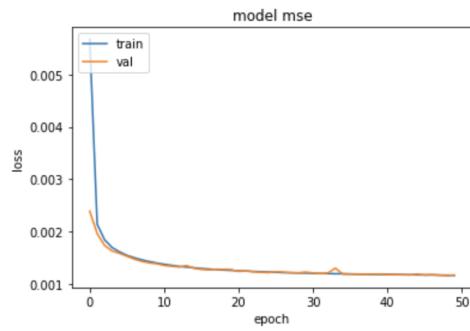
A continuación se adjunta una tabla en la cual se puede encontrar distintos parámetros característicos para cada modelo:

Gaussiano: Variance = 0.1				
Modelos	Num.Epocas	Batch_size	T.Entrenamiento (min)	R2 Score
Modelo 1	80	70	40	0.96
Modelo 2	120	70	46	0.92
Modelo 3	50	10	16	0.96
Modelo 4	40	5	17	0.96
Modelo 5: U-Net	100	25	42	0.05

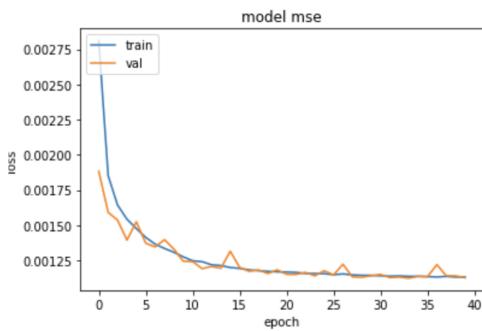
Plot Training and Validation Loss values:



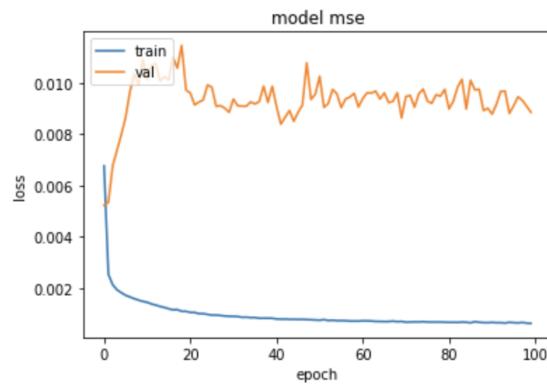
Modelo 3



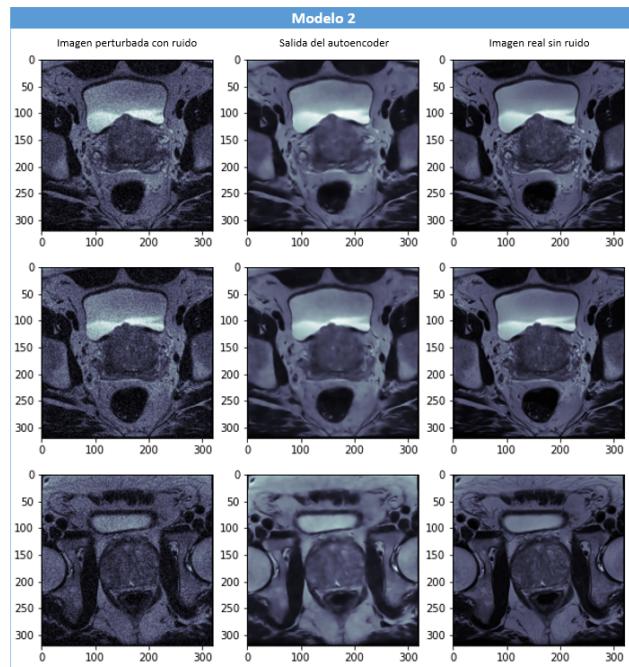
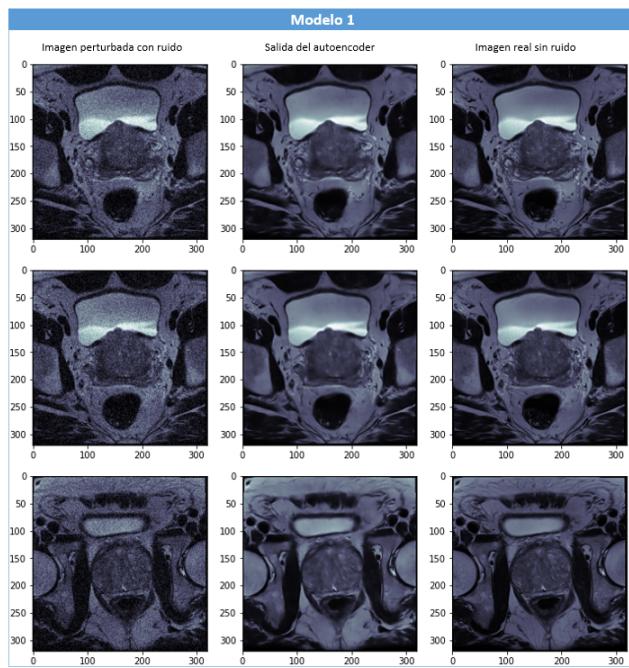
Modelo 4

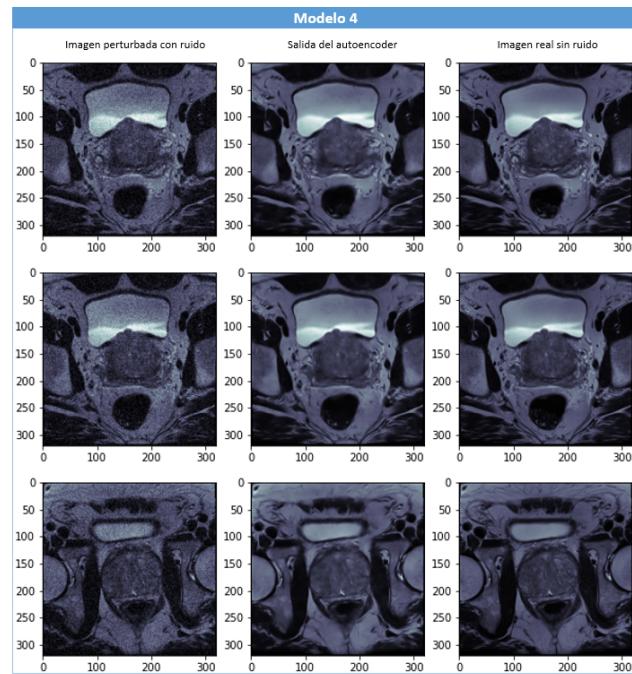
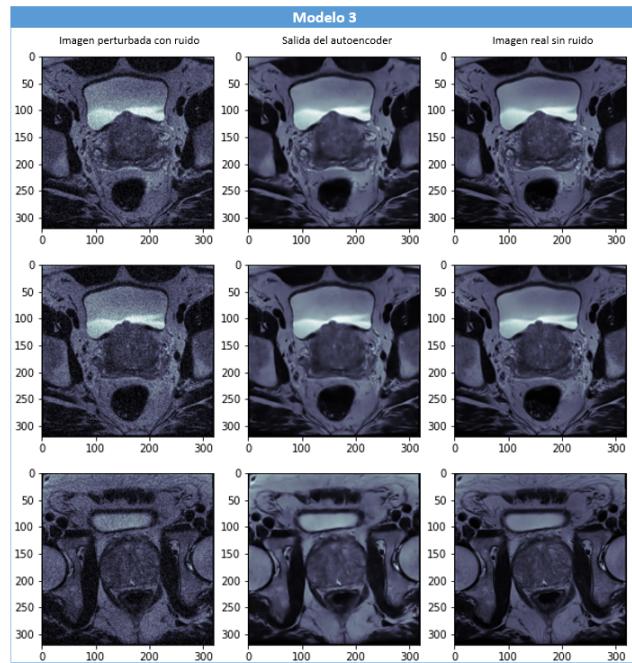


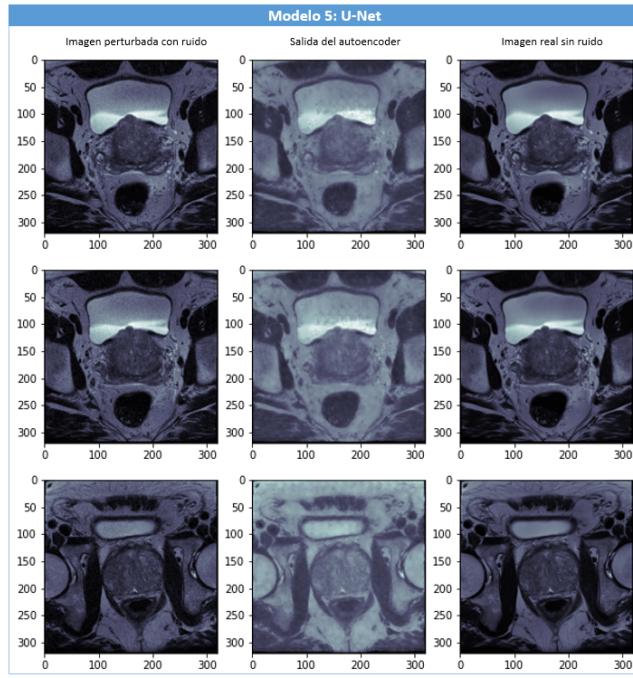
U-Net



Las imágenes obtenidas son:





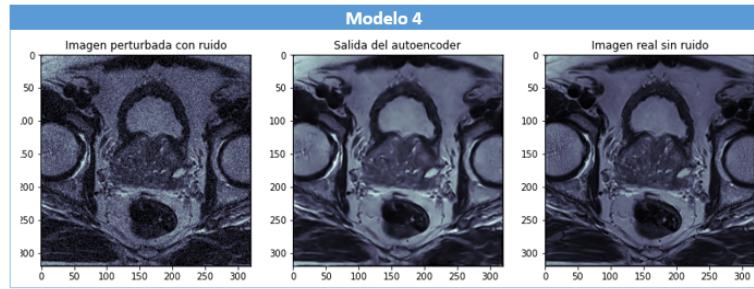
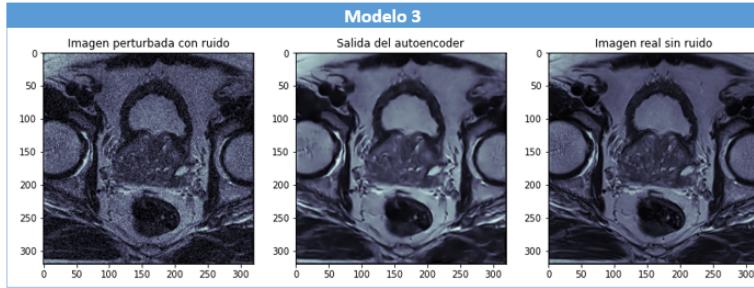
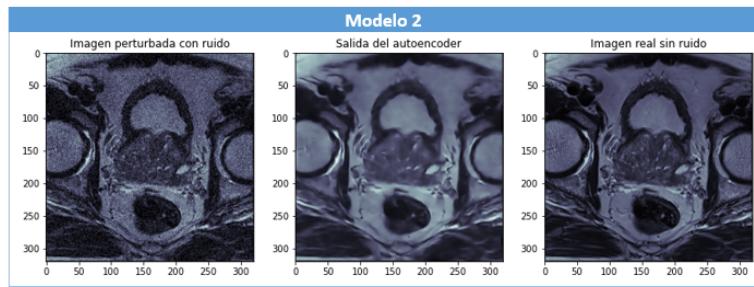
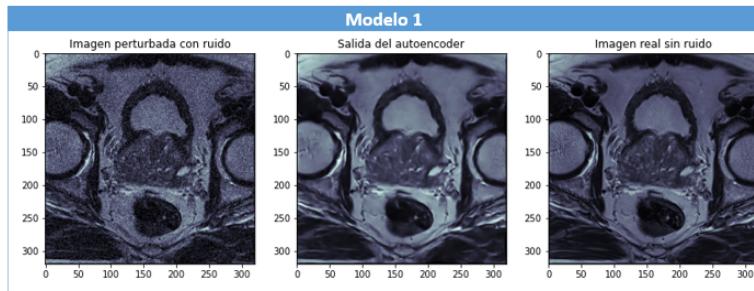


De nuevo se observa que el entrenamiento de los primeros cuatro modelos tienen resultados satisfactorios mientras que la red U-Net no es capaz de generar imágenes válidas que puedan ser utilizadas por un experto.

6.2.2.2. Validación de los Modelos

Los resultados obtenidos durante la validación son:

Modelos	R ² Entrenamiento	R ² Validación
Modelo 1	0.96	0.93
Modelo 2	0.92	0.78
Modelo 3	0.96	0.94
Modelo 4	0.96	0.94
U-Net	0.05	--



Tal y como ocurre en el modelo Gauss anterior, se produce una disminución insignificante en los valores de R2 excepto para el modelo 2, presentando una gran disminución en el valor de R2. Este hecho, nos puede indicar que durante el entrenamiento del modelo se ha producido sobreajuste.

6.3. Ruido Riciano

6.3.1. Primer Ruido Riciano

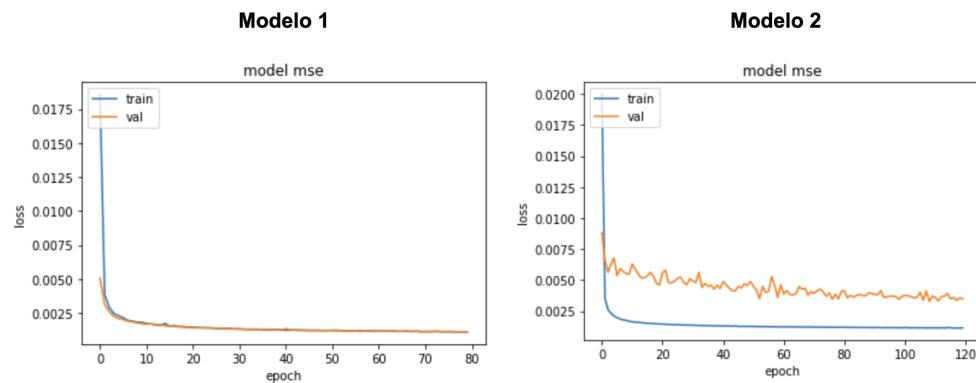
- Primer ruido Riciano: Noise-signal = 18; S0 = 1.5

6.3.1.1. Entrenamiento de los Modelos

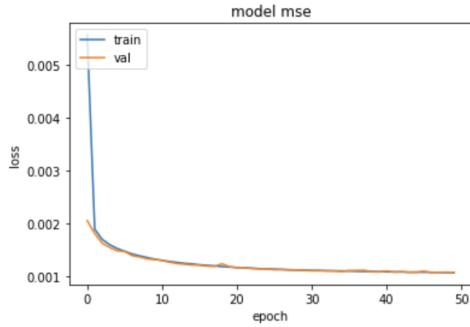
Los parámetros característicos para cada modelo son:

Riciano: Noise-signal 18; S0 = 1.5				
Modelos	Num.Epocas	Batch_size	T.Entrenamiento (min)	R2 Score
Modelo 1	80	70	37	0.97
Modelo 2	120	70	42	0.83
Modelo 3	50	10	16	0.97
Modelo 4	40	5	17	0.97
Modelo 5: U-Net	100	25	42	0.84

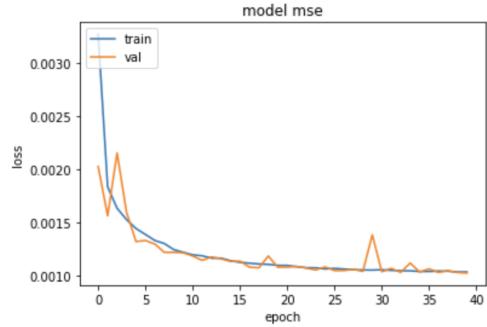
Las gráficas de entrenamiento obtenidas para cada modelo son:



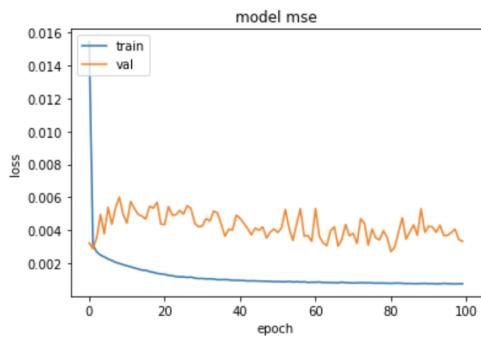
Modelo 3



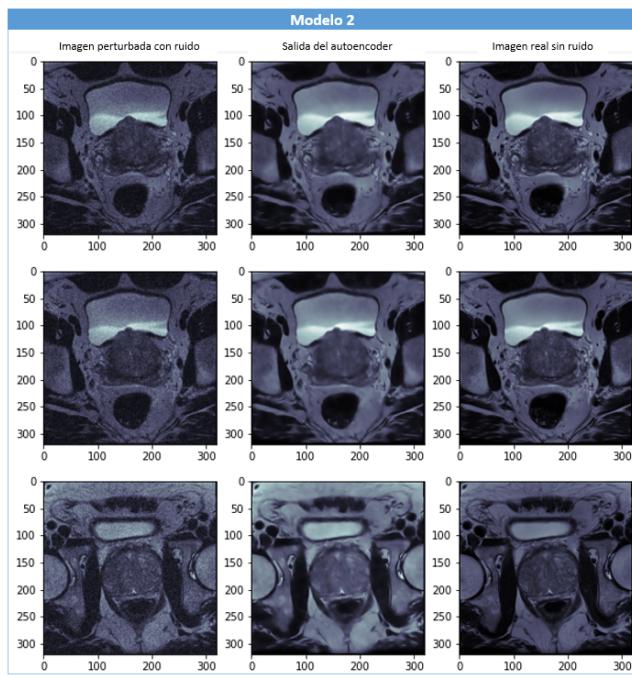
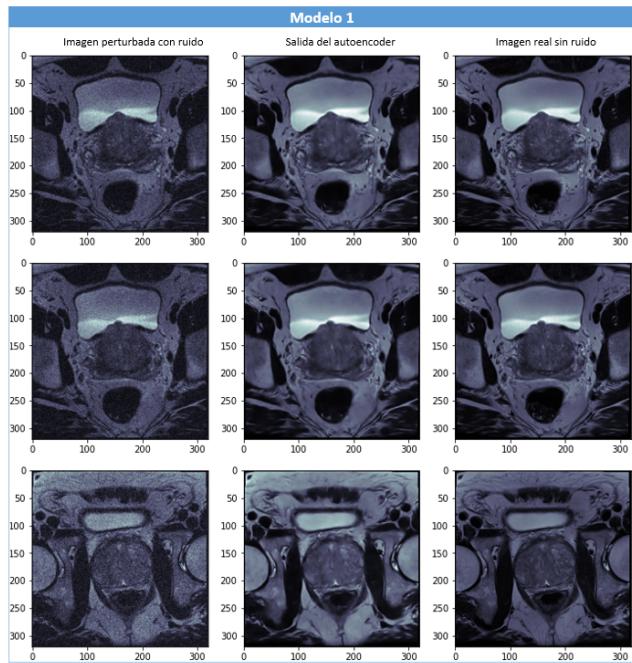
Modelo 4

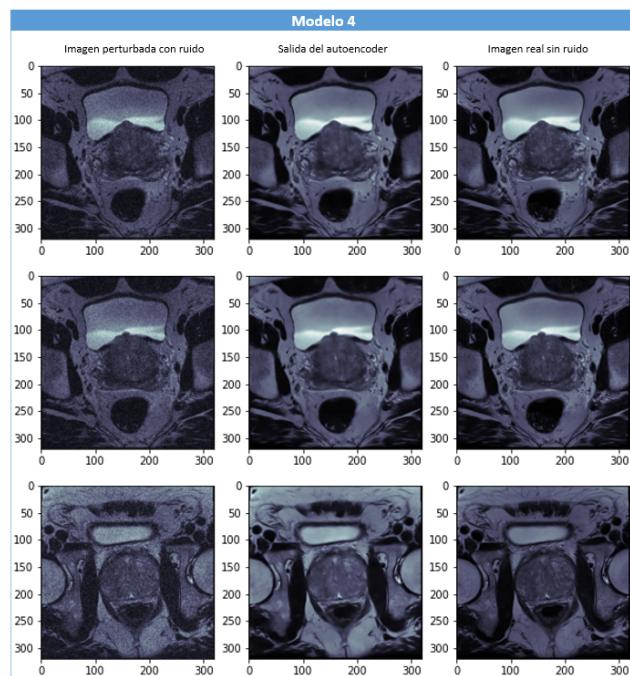
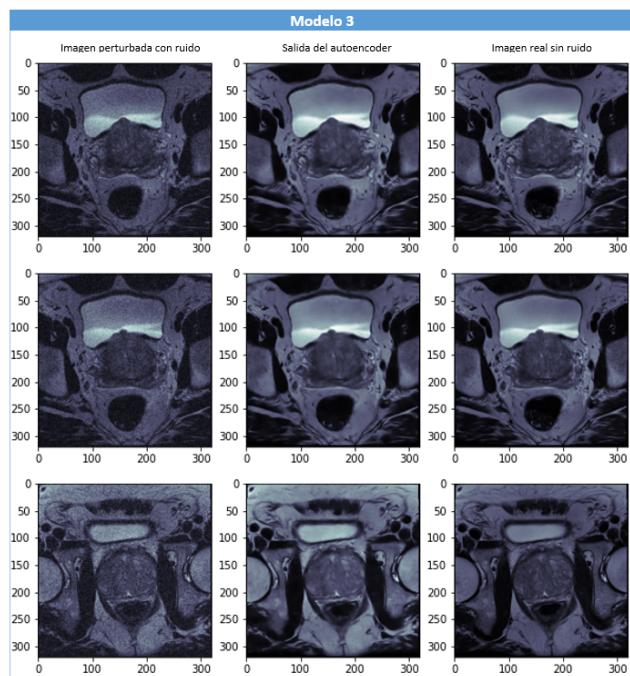


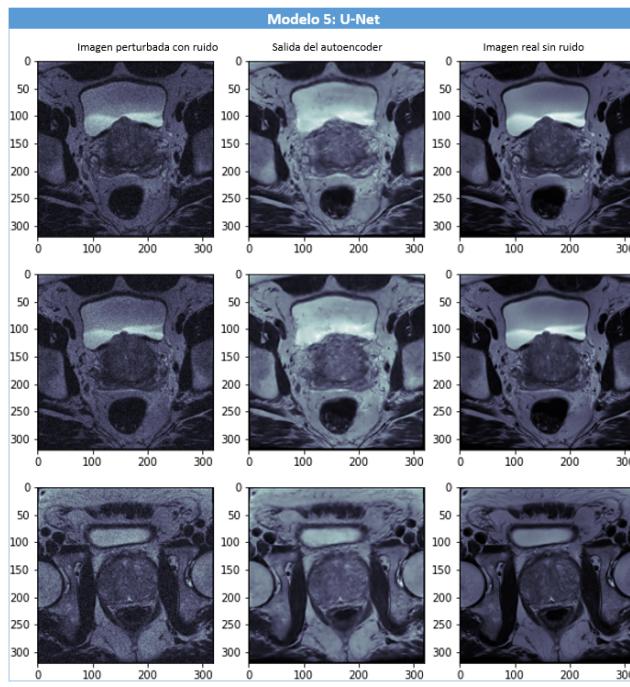
U-Net



Las imágenes obtenidas son:





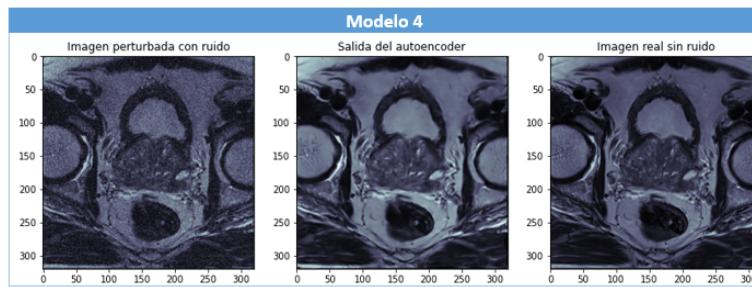
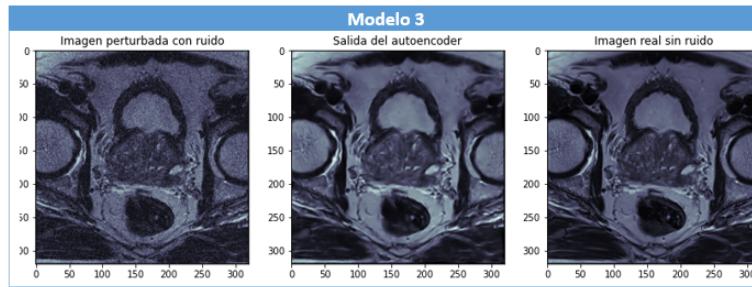
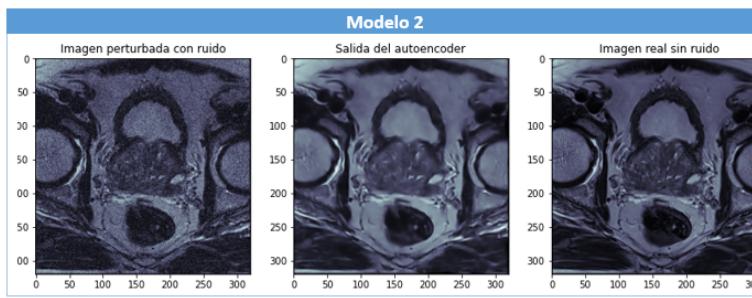
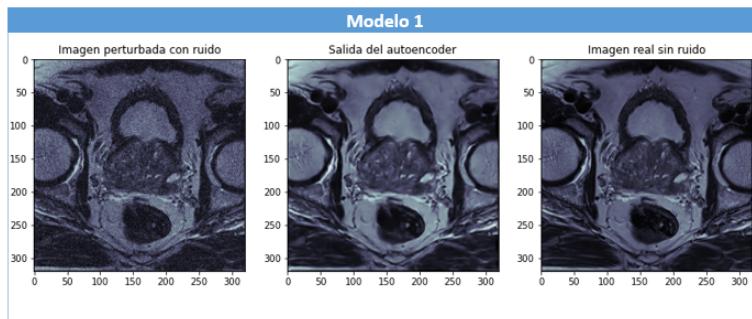


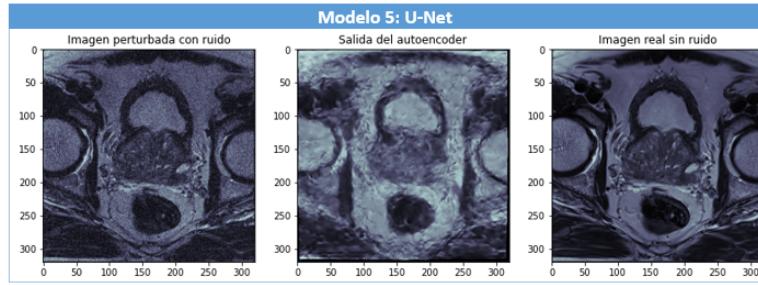
Durante esta fase de entrenamientos, los modelos 1,3,4 han dado resultados buenos (elevados valores de R²) mientras que el modelo 2 y la red U-Net a pesar de obtener un valor relativamente elevado de R², las imágenes de salida presentan una menor calidad (ligeramente borrosas).

6.3.1.2. Validación de los Modelos

Los resultados obtenidos durante la validación son:

Modelos	R ² Entrenamiento	R ² Validación
Modelo 1	0.97	0.94
Modelo 2	0.83	0.84
Modelo 3	0.97	0.94
Modelo 4	0.97	0.94
U-Net	0.84	0.37





Se puede observar que los modelos 1,3 y 4 tienen buenos resultados (R2 elevado y las imágenes de salida tienen una resolución óptima). Por el contrario los modelos 2 y red U-Net presentan valores menores de R2 así como imágenes de salida difuminadas.

6.3.2. Segundo Ruido Riciano

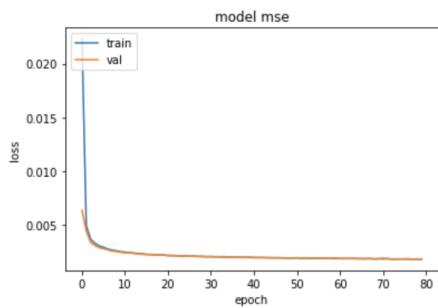
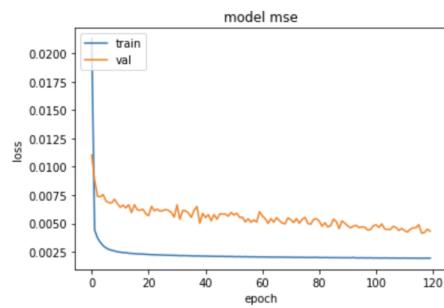
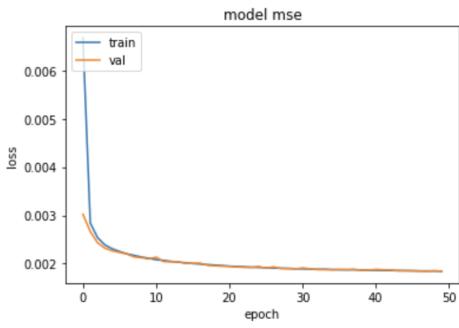
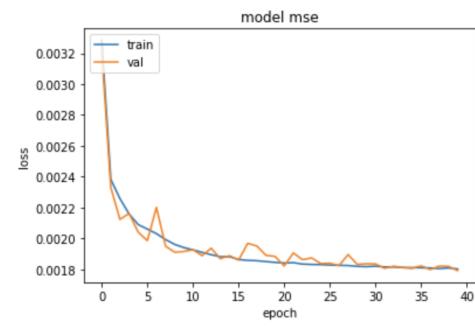
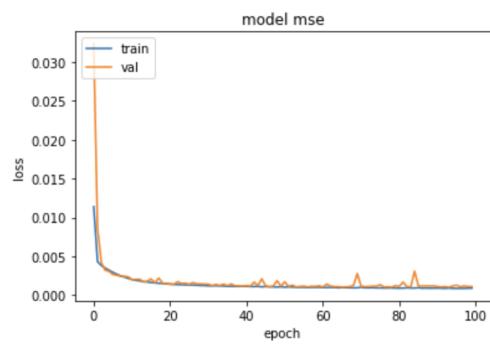
- Segundo ruido Riciano: Noise-signal = 14; S0 = 2

6.3.2.1. Entrenamiento de los Modelos

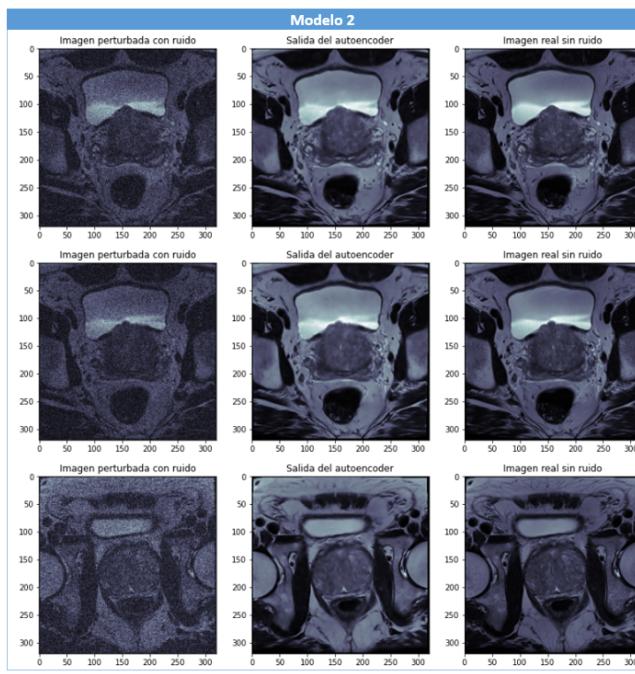
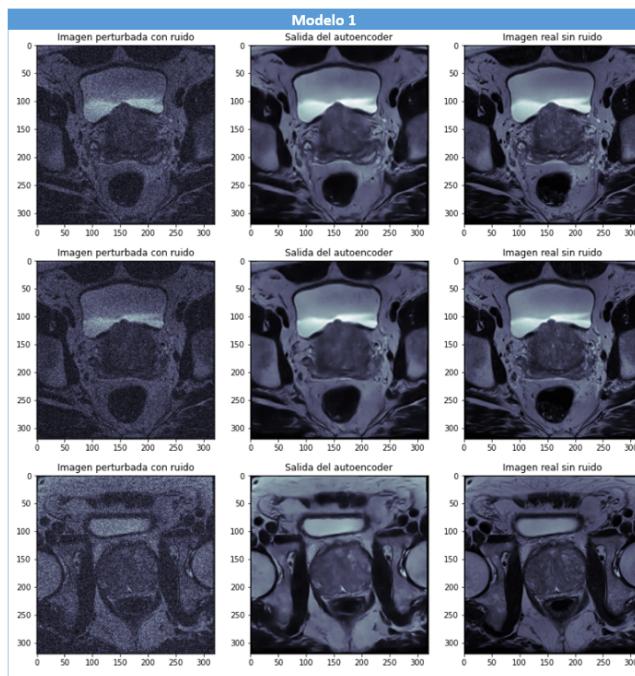
Los parámetros característicos para cada modelo son:

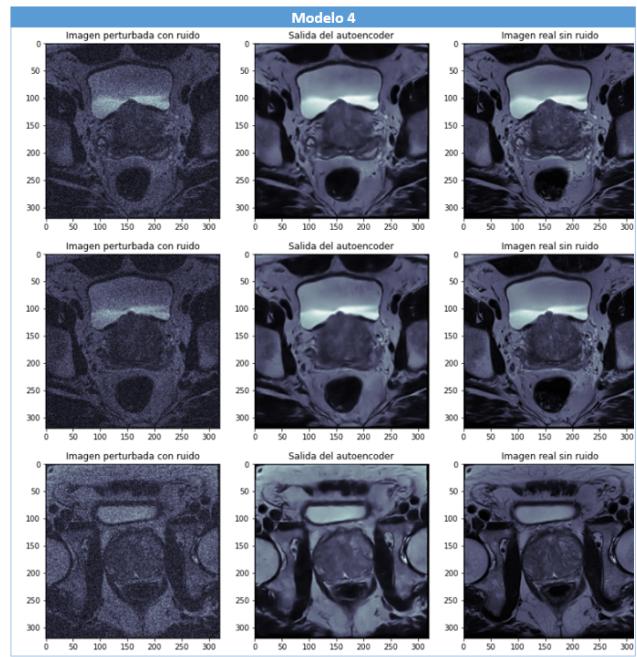
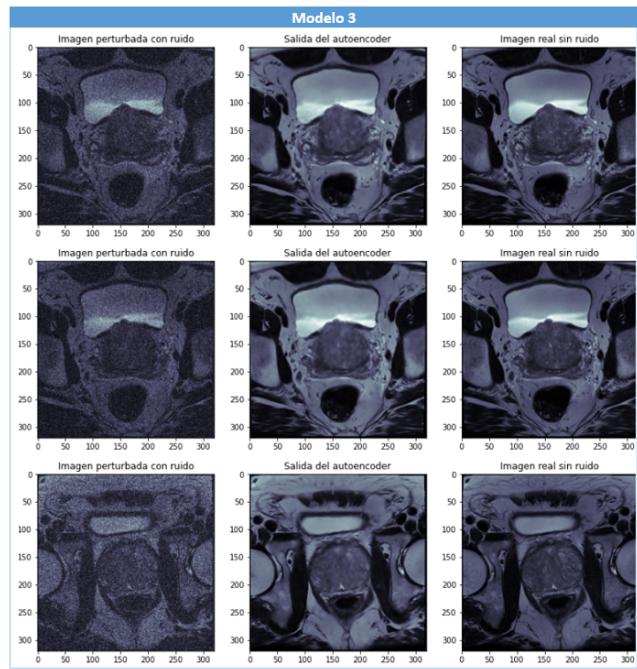
Riciano: Noise-signal 14; S0 = 2				
Modelos	Num.Epocas	Batch_size	T.Entrenamiento (min)	R2 Score
Modelo 1	80	70	40	0.94
Modelo 2	120	70	132	0.80
Modelo 3	50	10	47	0.94
Modelo 4	40	5	17	0.94
Modelo 5: U-Net	100	25	42	0.96

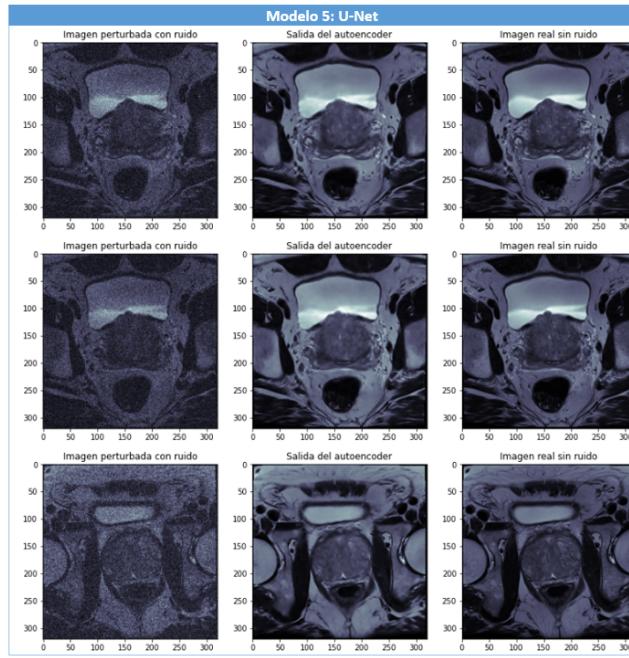
A continuación se presentan las correspondientes gráficas de entrenamiento:

Modelo 1**Modelo 2****Modelo 3****Modelo 4****U-Net**

Las imágenes obtenidas son:





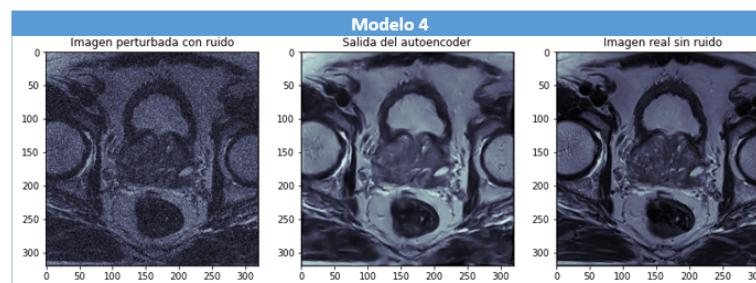
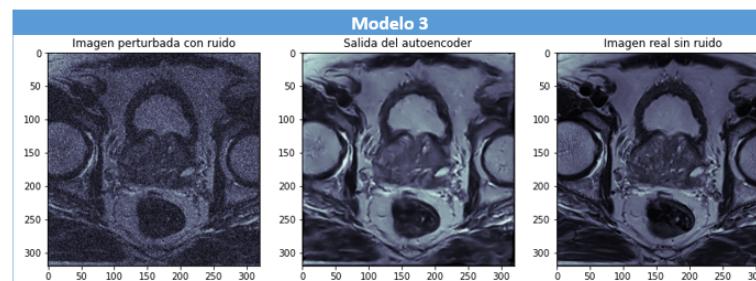
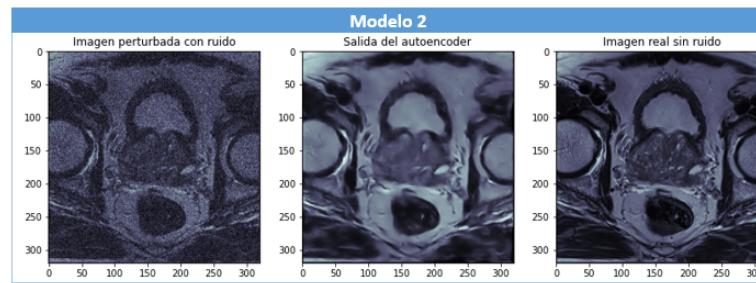
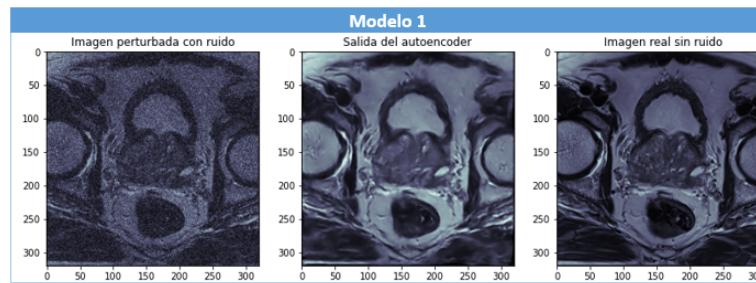


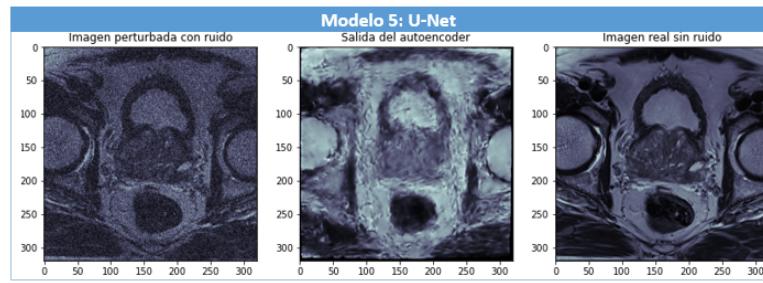
Durante esta fase de entrenamiento, podemos concluir que todos los los modelos ofrecen resultados positivos incluida la red U-Net.

6.3.2.2. Validación de los Modelos

Los resultados obtenidos durante la validación son:

Modelos	R ² Entrenamiento	R ² Validación
Modelo 1	0.94	0.84
Modelo 2	0.80	0.85
Modelo 3	0.94	0.85
Modelo 4	0.94	0.84
U-Net	0.96	0.42





En este caso se observa una mayor disminución en los valores de R2 con respecto a los obtenidos durante el entrenamiento. Debido al elevado grado de perturbación, las imágenes de salida pierden definición sobre todo en pequeños detalles.

Respecto a los resultados obtenidos utilizando la red U-Net, el modelo no ha sido capaz de reconstruir la imagen de forma adecuada, obteniendo imágenes distorsionadas.

7. Conclusiones

En este trabajo se han utilizado modelos capaces de realizar “*image denoising*” utilizando un conjunto de RMI de próstatas amplio en formato DICOM.

El desarrollo de este proyecto se ha enfocado en el tratamiento de imágenes perturbando las con diferentes tipos y rangos de ruidos.

Una vez presentados los resultados, se pueden obtener las siguientes conclusiones:

- Los datos de entrada se han normalizado según la fórmula:

$$X \text{ normalizado} = \frac{X - X \text{ mín}}{X \text{ máx} - X \text{ mín}}$$

A pesar de obtener resultados favorables, esta técnica presenta un problema:

Los datos de entrada se comprimen entre unos límites empíricos (el máximo y el mínimo de píxeles), por tanto, si existe ruido, este se amplía.

- A medida que se aumenta la densidad ruido, los modelos (autoencoders) pierden su capacidad de reconstrucción de la imagen, obteniéndose imágenes poco definidas, es decir, presentan cierto grado de borrosidad (se ven difuminadas y se pierden detalle).
- La naturaleza del tipo de ruido afecta a la capacidad de aprendizaje de los modelos. Un ejemplo evidente de esto, es el caso del ruido Riciano, en el cual los resultados obtenidos fueron peores que los resultados de los ruidos Gaussiano y Salt & Pepper.
- En el Modelo 2 una de las capas es Dropout. Inicialmente se pensó que esta capa nos ayudaría a obtener mejores resultados ya que evita el overfitting, sin embargo, los resultados obtenidos con este modelo han sido desfavorables.
- Para evitar *overfitting* y realizar entrenamientos más rápidos, en el Modelo 4 se introdujo una capa de Batch normalization. Los resultados obtenidos utilizando esta técnica de normalización fueron positivos.
- La red U-Net es muy utilizada en proyectos de segmentación en imágenes médicas. Sin embargo, debido a la naturaleza de su estructura (tiene una gran capacidad de

compresión), en proyectos enfocados a la reducción de ruido no se obtienen buenos resultados.

- Trabajar con imágenes requiere gran capacidad de computación lo que nos ha supuesto realizar el entrenamiento de las redes utilizando la GPU de Google Colab.
- Los tiempos de entrenamiento de los modelos no han sido elevados. Además se ha observado que la naturaleza de ruido no influye considerablemente en el tiempo de entrenamiento.
- A pesar de los buenos resultados obtenidos utilizando los modelos 1,3 y 4, sería necesario que un experto realizase la validación de manera manual de las imágenes resultantes. Por ese motivo el uso clínico de estos modelos se limita a una herramienta auxiliar.

8. Trabajo Futuro

Como trabajo futuro para este proyecto podemos destacar:

- Utilización de la librería Fast.ai sobre el framework Pytorch en lugar de utilizar Keras sobre TensorFlow. Con ello se podría ahorrar tiempo en la toma de decisiones/parámetros en los modelos.
- Realizar modelos con una mayor gama de intensidades de ruido.
- Realizar un modelo con una estructura única que sea capaz de reducir cualquier tipo de ruido, independientemente de la densidad y naturaleza del mismo.
- Construir una interfaz de usuario donde se pueda realizar la reducción de ruido y el experto pueda ir validando los resultados obtenidos.
- Implementar técnicas de “*Denoising*“ a través de Transfer Learning.
- Realizar el presente proyecto utilizando imágenes de resonancia obtenidas de otras partes del cuerpo humano como cerebro, próstata, columna vertebral etc.
- Eliminación de ruido en imágenes obtenidas mediante otras técnicas (rayos-x, ecografía, etc).

9. Bibliografía

- [1] Statista. Número de exámenes por resonancia magnética (IRM) realizados en los países seleccionados en 2007.
- [2] MARTINI, G. (2019, junio 21). La lista d'attesa è lunga mesi ma pagando la visita è domani. Recuperado 19 de mayo de 2020, de <https://www.lastampa.it/cronaca/2016/09/19/news/la-lista-d-attesa-e-lunga-mesi-ma-pagando-la-visita-e-domani-1.34807211>
- [3] Useros, V. (2019, octubre 21). La demora en resonancias magnéticas, mayor en los hospitales públicos que en los centros concertados. Recuperado el 19 de mayo de 2019. <https://www.elmundo.es/comunidad-valenciana/2019/10/21/5da9aab3fdddf436c8b465d.html>
- [4] Kumar, K. K. (2020). A NOVEL CONVOLUTIONAL AUTOENCODERS BASED MULTILEVEL DENOISING FRAMEWORK FOR MEDICAL IMAGE DATA. *Journal of Critical Reviews*, 7(2), 30-37.
- [5] Anwer, D. A. (2020). The Impact of Neural Network Techniques in the Optimization of the Image Processing. *Journal of Southwest Jiaotong University*, 55(2).
- [6] Gondara, L. (2016, December). Medical image denoising using convolutional denoising autoencoders. In *2016 IEEE 16th International Conference on Data Mining Workshops (ICDMW)* (pp.241-246).
- [7] Nowak, R. D. (1999). Wavelet-based Rician noise removal for magnetic resonance imaging. *Transactions on Image Processing*, 8(10), 1408-1419.
- [8] Cárdenas-Blanco, A., Tejos, C., Irarrazaval, P., & Cameron, I. (2008). Noise in magnitude of magnetic resonance images. *Concepts in Magnetic Resonance Part A: An Educational Journal*, 32(6), 409-416.

Isa, I. S., Sulaiman, S. N., Mustapha, M., & Darus, S. (2015). Evaluating denoising performances of fundamental filters for t2-weighted MRI images. Procedia Computer Science, 60, 760-768.

[10] Rayos X. (s. f.). Recuperado el 10 de enero de 2020, de https://www.xtal.iqfr.csic.es/Cristalografia/parte_02.html

[11] Pardell, J. (s. f.). Física de los Rayos X. Recuperado el 10 de enero de 2020, de <https://www.pardell.es/fisica-de-los-rayos-x.html>

[12] Calzado, A., & Geleijns, J. (1). Tomografía computarizada. Evolución, principios técnicos y aplicaciones. Revista Española De Física Médica, 11(3). Recuperado a partir de <https://revistadefisicamedica.es/index.php/rfm/article/view/115>

[13] Burbano Arrázola, M. B. (2012, noviembre). Tomografía por Emisión de Positrones: principios y aplicaciones.

[14] Sánchez-Salmón, A. S., & Ruibal, A. R. (2006). Valor actual de la Tomografía por Emisión de Positrones durante el seguimiento del Cáncer de Próstata. Recuperado de <http://scielo.isciii.es/pdf/urol/v59n10/original8.pdf>

[15] Escuela Universitaria de Tecnología Médica Udelar, Montevideo, Uruguay Comité de Tecnólogos de ALASBIMN, & Núñez , M. N. (2008). Tomografía por emisión de positrones (PET): Fundamentos. Recuperado de https://www.alasbimn.net/comites/tecnologos/material/Fundamentos_PET.pdf

[16] Fernández-Friera, L. F. F., García-Álvarez, A. G., & Ibáñez, B. I. (2013). Imaginando el futuro del diagnóstico por imagen. Revista Española de Cardiología, 66(2), 134-143. <https://doi.org/10.1016/j.recesp.2012.10.012>

[17] A.Kaufmann, P. A. K., & F.Di Carli, M. F. D. C. (2009). Hybrid SPECT/CT and PET/CT Imaging: The Next Step in Noninvasive Cardiac Imaging. ScienceDirect, 39(5), 341-347. <https://doi.org/10.1053/j.semnuclmed.2009.03.007>

[18] Ministerio de Salud Pública del Ecuador Instituto Ecuatoriano de Seguridad Social Sociedad de Lucha Contra el Cáncer. (2012). NORMATIVA DE APLICACIÓN DE PET-CT.

- ASCIRES, Grupo Biomédico ASCIRES. (2018). CETIR y GE Healthcare presentan el PET/RM, la última tecnología en Diagnóstico por Imagen [Diapositivas]. Recuperado de <https://www.cetir.com/wp-content/uploads/2018/03/PET-RM-CETIR-Grupo-ASCIRES.pdf>
- [20] Bibliografías y vidas. La enciclopedia bibliográfica en línea. (2004). Recuperado el 3 de febrero de 2020, de <https://www.biografiasyvidas.com/>
- [21] Canals, Mauricio. (2008). HISTORIA DE LA RESONANCIA MAGNÉTICA DE FOURIER A LAUTERBUR Y MANSFIELD: EN CIENCIAS, NADIE SABE PARA QUIEN TRABAJA. Revista chilena de radiología. 14. 10.4067/S0717-93082008000100009.
- [22] Universidad de Santiago de Compostela. (s. f.). Resonancia Magnética . Recuperado el 3 de marzo de 2020, de <https://www.usc.gal/gl/investigacion/riaidt/rm/rmn/introducion.html>
- [23] Merrit CRB. Physics of ultrasound. En: Rumack CM, Wilson SR, Charboneau JW. Diagnostic Ultrasound. 2nd Ed. Mosby.1997; 1: 19-33.
- [24] Hendee WR. Reflection Imaging (ultrasound) In: Putman CE, Ravin CE. Diagnostic Imaging. 1st Ed. Philadelphia Saunders. 1988; 1: 62-69.
- [25] Cabrera R, Alba F. Historia de la Ultrasonografía en Chile: Ecografía en Ginecología. Rev Chil Ultrasonog 2002; 5: 27-34.
- [26] Pineda, V. C., Macías, P. M., y Bernal, G. A., (2012). Principios Físicos básicos del ultrasonido. Investigación en Discapacidad, 1(1), 25-34.
- [27] Rojas, G. M., Ardilla, T. G., Jaramillo, B. N, (2017). Artificios ecográficos más frecuentes del modo B: Consideraciones para la adecuada interpretación en la práctica del ultrasonido aplicado a los cuidados médicos perioperatorios. Anestesia en México, 29 (3), 4-15.
- [28] SEINAP, (2012). Revista Española de Pediatría clínica e Investigación, 68 (2), 1-153.
- [29] Asociación Nacional de Fabricantes Eléctricos. Publicaciones varias. Disponible en: DicomStandard.org
- [30] Wiki oficial de GDCM.

- [31] <https://steemit.com/spanish/@eniolw/dicom-estudio-conceptual-del-estandar>
- [32] OLEG S. PIANYKH, Digital Imaging and Communications in Medicine (DICOM) A Practical Introduction and survival Guide.
- [33] <https://brainder.org/2012/09/23/the-nifti-file-format/>
- [34] http://scielo.sld.cu/scielo.php?script=sci_arttext&pid=S1684-18592014000100008
- [35] https://wiki.hl7.org/Introduction_to_HL7:_Content#What_is_HL7.3F
- [36] Dolin RH, et al. HL7 Clinical Document Architecture. J Am Med Inform Assoc. 2001 Nov-Dec;8(6):552-69.
- [37] GUERRERO, T. G., & LUCIO, C. G. L. (2018). INTELIGENCIA ARTIFICIAL: UN NUEVO ALIADO DE LOS MÉDICOS. Recuperado de <https://lab.elmundo.es/inteligencia-artificial/salud.html>
- [38] Pérez, I. P. [StarsConf]. (2019, enero 18). Deep Learning en Medicina [Archivo de vídeo]. Recuperado de <https://www.youtube.com/watch?v=wSew-KfDl4M&t=696s>
- [39] Sáenz Bajo, N., Barrios Rueda, E., Conde Gómez, M., Domínguez Macías, I., López Carabaño, A., & Méndez Díez, C. (2002). Uso de redes neuronales en medicina: a propósito de la patología dispéptica. Elsevier, 30(2), 99-102. Recuperado de <https://www.elsevier.es/es-revista-atencion-primaria-27-articulo-uso-redes-neuronales-medicina>
- [40] Rodrigues, W. (2018). SineReLU — An Alternative to the ReLU Activation Function. Recuperado de <https://medium.com/@wilder.rodrigues/sinerelu-an-alternative-to-the-relu>
- [41] Curiously. (2019, julio). Build a simple Neural Network with TensorFlow.js | Deep Learning for JavaScript Hackers (Part III). Recuperado de <https://www.curiously.com/posts/build-a-simple-neural-network-with-tensorflow-js/>
- [42] Ronneberger.O, Fischer.F, Thomas.B. (2015, mayo 18). U-Net: Convolutional Networks for Biomedical Image Segmentation. Recuperado el 20 de febrero de 2020, de <https://arxiv.org/pdf/1505.04597.pdf>

- Sharma, A. (2018, abril 4). Implementing Autoencoders in Keras: Tutorial. Recuperado el 20 de febrero de 2020, de <https://www.datacamp.com/community/tutorials/autoencoder-keras-tutorial>
- [44] Alqadi, Ziad. (2018). Salt and Pepper Noise: Effects and Removal. International Journal on Electrical Engineering and Informatics. 2.
- [45] Abd Al-salam Selami, Ameen & Fadhil, Ahmed. (2016). A Study of the Effects of Gaussian Noise on Image Features. Kirkuk University Journal / Scientific Studies (1992-0849). 11. 152 - 169.
- [46] Gudbjartsson, H., & Patz, S. (1995). The Rician distribution of noisy MRI data. Magnetic resonance in medicine, 34(6), 910–914. <https://doi.org/10.1002/mrm.1910340618>
- [47] Barbará Morales, Eduardo & Bao, Reinier & González, Evelio. (2012). Comparación de Algoritmos de Segmentación de Ruido Aplicados a Imágenes de Resonancia Magnética. Revista Electrónica, Automática y Comunicaciones. 33. 8-18. 10.1234/rielac.v33i3.134.
- [48] Janeth , S., & Ordóñez, X. G. (2014). Estadística descriptiva e inferencial. Madrid, España: Udima.