

Detección de neumonía a partir de radiografías de tórax (CXR)

Mario Ubierna San Mamés

Máster universitario en Ciencia de Datos - Universitat Oberta de Catalunya
Área de la medicina

Jordi de la Torre Gallart

06/2022



Esta obra está sujeta a una licencia de Reconocimiento-NoComercial-SinObraDerivada [3.0 España de Creative Commons](https://creativecommons.org/licenses/by-nc-nd/3.0/es/)

FICHA DEL TRABAJO FINAL

Título del trabajo:	<i>Detección de neumonía a partir de radiografías de tórax (CXR)</i>
Nombre del autor:	<i>Mario Ubierna San Mamés</i>
Nombre del consultor/a:	<i>Jordi de la Torre Gallart</i>
Nombre del PRA:	<i>Jordi de la Torre Gallart</i>
Fecha de entrega (mm/aaaa):	06/2022
Titulación:	<i>Máster universitario en Ciencia de Datos</i>
Área del Trabajo Final:	<i>Área Medicina (TFM-Med)</i>
Idioma del trabajo:	<i>Castellano</i>
Palabras clave	<i>Redes neuronales convolucionales, radiografía de tórax, neumonía.</i>
Resumen del Trabajo (máximo 250 palabras): <i>Con la finalidad, contexto de aplicación, metodología, resultados i conclusiones del trabajo.</i>	
<p>La finalidad de este trabajo es la construcción de un modelo de <i>deep learning</i>, que nos permita determinar si a partir de radiografías de tórax (CXR) un paciente presenta una neumonía.</p> <p>Aunque las redes neuronales artificiales aparecen por primera vez en 1958, cuando apareció el perceptrón, se han ido mejorando las diferentes propuestas hasta tal punto que desde hace 10 años hasta día de hoy son imprescindibles en nuestra sociedad gracias al elevado rendimiento que proporcionan. Dentro de las redes neuronales, las redes neuronales convolucionales (CNN) han adquirido una gran importancia en los últimos años, debido al alto nivel de precisión que son capaces de alcanzar en tareas de clasificación de imágenes.</p> <p>Para poder lograr el objetivo del proyecto, se ha hecho uso de un elevado número de radiografías de tórax, éstas se han preprocesado y posteriormente se han introducido a la red neuronal convolucional con el objetivo de clasificar las mismas.</p> <p>Gracias a las más de 25000 radiografías y el modelo generado se ha conseguido determinar si un paciente padece una neumonía con una elevada precisión, permitiendo así dar un pequeño paso en el avance biotecnológico.</p>	

Abstract (in English, 250 words or less):

The purpose of this work is the construction of a deep learning model, which allows us to determine if a patient has pneumonia from chest radiographs (CXR).

Although artificial neural networks appear for the first time in 1958, when the perceptron appeared, the different proposals have been improving to such an extent that from 10 years ago to the present day they are essential in our society thanks to the high performance they provide. Within neural networks, convolutional neural networks (CNN) have gained great importance in recent years, due to the high level of precision they are capable of achieving in image classification tasks.

To achieve the objective of the project, a large number of chest radiographs have been used, these have been preprocessed and subsequently entered into the convolutional neural network to be able to classify them.

Thanks to the more than 25,000 chest radiographs and the generated model, it has been possible to determine with great precision if a patient suffers from pneumonia, thus allowing a small step in biotechnological progress to be taken.

Índice

1. Introducción.....	1
1.1 Contexto y justificación del Trabajo.....	1
1.2 Motivación personal.....	2
1.3 Objetivos del Trabajo.....	3
1.4 Enfoque y método seguido.....	3
1.5 Planificación del Trabajo	5
1.6 Breve resumen de productos obtenidos	7
1.7 Breve descripción de los otros capítulos de la memoria.....	7
2. Estado del arte	8
2.1 Marco teórico sobre la neumonía	8
2.2 Marco teórico sobre las redes neuronales convolucionales	10
2.3 Estudios similares	12
3. Desarrollo	15
3.1 Extracción y compresión de los datos	15
3.1.1 Dónde y cómo se obtuvieron los datos	15
3.1.2 Estructura del repositorio	17
3.1.3 Estructura de los datos	19
3.2 Preprocesado de los datos.....	21
3.2.1 Lectura y filtrado de los datos	21
3.2.2 Creación de los conjuntos de entrenamiento, validación y test.....	22
3.2.3 Transformación de las imágenes	22
3.2.3 Balanceo de los datos.....	23
3.2.4 Generación de los ficheros de salida	24
3.3 Creación de los modelos	25
3.3.1 Data Generator	25
3.3.2 Carga de las imágenes preprocesadas	26
3.3.3 Creación de las redes neuronales convoluciones con transfer learning.....	26
3.3.4 Fine-tuning	28
4. Experimentos y resultados	30
5. Conclusiones.....	31
6. Líneas de trabajo futuro	33
7. Glosario	34
8. Bibliografía	35
9. Anexos	38

Índice de figuras

Ilustración 1 - (a) radiografía paciente normal, (b) radiografía paciente con neumonía.	2
Ilustración 2 - Cross Industry Standard Process for Data Mining (CRISP-DM). .	4
Ilustración 3 - Planificación del proyecto.	6
Ilustración 4 - Marco teórico sobre la red neuronal convolucional (RNC).....	11
Ilustración 5 - Número de publicaciones en PubMed con la consulta "classification deep convolutional neural networks".	12
Ilustración 6 - "RSNA Pneumonia Detection Challenge" repositorio de <i>Kaggle</i>	15
Ilustración 7 - Comando de descarga del repositorio.	16
Ilustración 8 - Descarga manual de los ficheros del repositorio (uno a uno). ...	16
Ilustración 9 - Descarga manual de los ficheros del repositorio (todos a la vez).	16
Ilustración 10 - <i>Dashboard</i> repositorio de <i>Kaggle</i>	17
Ilustración 11 - Estructura del repositorio de <i>Kaggle</i>	17
Ilustración 12 - Estructura del repositorio de <i>GitHub</i>	18
Ilustración 13 - Estructura de "stage_2_detailed_class_info.csv"	20
Ilustración 14 - Estructura de "stage_2_train_labels.csv"	20
Ilustración 15 - Lectura de los datos.....	21
Ilustración 16 - Imagen original vs imagen aumentada.	24
Ilustración 17 - Capas de un modelo <i>EfficientNet</i>	27

Índice de tablas

Tabla 1 - Tamaños de las imágenes para cada red neuronal.	23
Tabla 2 - Comparativa entre el nombre de la imagen original y las aumentadas.	24

1. Introducción

1.1 Contexto y justificación del Trabajo

La neumonía [1] es una infección del sistema respiratorio que afecta a los pulmones, es decir, puede generar una inflamación en los bronquiolos, bronquios y alvéolos pulmonares debido a microorganismos como hongos, bacterias y virus [2].

Aunque el término de neumonía está al día entre nosotros y sobre todo después del SARS-CoV2 [3], no somos conscientes de la importancia de esta enfermedad. Según Rosario Menéndez, neumóloga de la Sociedad Española de Neumología y Cirugía Torácica (SEPAR) [4], cito textualmente: *“La neumonía es la primera causa de muerte por infección, y no está reconocida como la enfermedad peligrosa que es ni por la comorbilidad que supone”*.

Indagando aún más en la problemática, la Organización Mundial de la Salud (OMS) informa que el 15% de las defunciones entre los niños menores de 5 años están causadas por esta enfermedad [5], por lo tanto, este grupo de edad junto con las personas mayores de 60 años son los más afectados, debido al bajo sistema inmune que se presenta a esas edades.

A nivel nacional la problemática continúa, ya que solo en el año 2017 se ingresaron a 120000 personas dejando a su paso más de 10000 muertes [4], entrando España en el top de los países de la Unión Europea con más fallecidos por neumonía.

Entendiendo cómo de importante es esta enfermedad y las consecuencias de la misma, se ve de forma clara que es de vital importancia poder detectar y diagnosticar cuanto antes la neumonía, con el objetivo de salvar vidas.

Hasta el día de hoy, el diagnóstico de enfermedades pulmonares se puede hacer a partir de radiografías, resonancias magnéticas, broncoscopias, pruebas de esfuerzo... Todas ellas se complementan, pero por norma general se suele hacer uso de radiografías del tórax, ya que éstas junto con la opinión de un experto es suficiente para diagnosticar si existe la enfermedad y en qué grado. En la siguiente imagen se puede observar la diferencia entre un paciente sano y uno con neumonía a partir de una radiografía de tórax:

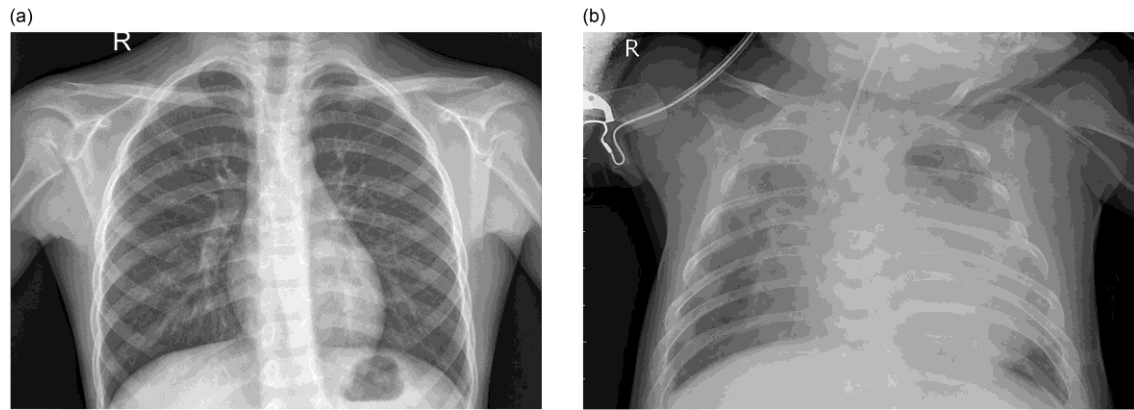


Ilustración 1 - (a) radiografía paciente normal, (b) radiografía paciente con neumonía [6].

El inconveniente en el diagnóstico de la neumonía es que se realiza de forma manual, en otras palabras, se necesita que un médico analice la radiografía para poder determinar si una persona padece de dicha enfermedad o no.

Por otro lado, hay que tener en cuenta el costo económico que supone, según la Organización Mundial de la Salud (OMS) el costo de antibióticos y pruebas necesarias para el diagnóstico de la neumonía solamente en niños supondría invertir 109 millones de dólares al año [5].

Ante este contexto surge la idea de este proyecto, el hacer uso del *deep learning*, y más concretamente de redes neuronales convolucionales (CNN), para poder diagnosticar si una persona padece neumonía o no sin la necesidad de la intervención médica, consiguiendo así reducir la saturación del sistema sanitario, además de reducir los costes al ser capaces de identificar la neumonía con antelación.

1.2 Motivación personal

El motivo principal por el que escogí este trabajo es básicamente por el interés personal en el ámbito sanitario. Considero que a día de hoy, existen muchas enfermedades tanto en países desarrollados como no desarrollados, en los que hacer uso de la tecnología permitirá resolver de una forma más eficiente el diagnóstico de enfermedades, consiguiendo así mejorar la calidad de vida de la sociedad.

Debido a ese motivo me parece de vital importancia proyectos como este o similares, en los que se busca investigar cómo afecta una enfermedad y a partir de redes neuronales poder diagnosticar si una persona padece neumonía o no, consiguiendo así reducir tiempos, o dicho en otras palabras, reducir el riesgo de mortalidad que conlleva una enfermedad.

Por otro lado, considero que el conocimiento que voy a obtener desarrollando este proyecto me va a aportar grandes cualidades para el mundo laboral, principalmente porque el *deep learning* está visible en una gran variedad de proyectos.

1.3 Objetivos del Trabajo

La hipótesis u objetivo principal de este trabajo es la generación de un modelo que permita determinar si una persona padece neumonía, a partir de radiografías de tórax.

Con el fin de alcanzar dicho objetivo, se plantean los siguientes objetivos secundarios:

- Entender la problemática de la neumonía, qué la causa, cuáles son las consecuencias, cómo se diagnostica a día de hoy, y el estado del arte en esta área.
- Obtener los datos del repositorio en *Kaggle* [7], entendiendo por datos las imágenes médicas en formato *DCM* [8].
- Realizar un preprocesado, para poder enviar a la red neuronal los datos que espera y mejorar la precisión de la misma.
- Creación de la red neuronal convolucional, encargada de recibir la imagen como entrada y capaz de clasificar si la persona padece neumonía.
- Evaluación del modelo, determinar cómo de bueno es para dar por alcanzado el objetivo principal.
- Búsqueda de los hiperparámetros idóneos, con el fin de mejorar la solución.
- Generar la documentación técnica del proyecto.
- Presentar y defender el proyecto ante un tribunal.

1.4 Enfoque y método seguido

A partir de las radiografías de tórax se quiere obtener un conocimiento para determinar si una persona tiene neumonía, para ello se va a hacer uso de una red neuronal convolucional, que use la imagen como entrada y nos devuelva si hay presencia de dicha enfermedad.

Tal y como se puede apreciar en el objetivo del proyecto, estamos en un caso de minería de datos. Por lo tanto, se va a hacer uso de la metodología *Cross Industry Standard Process for Data Mining (CRISP-DM)* [9], ya que ésta es una metodología estándar y es utilizada en un gran número de proyectos de minería de datos.

Cuando hablamos de metodología nos referimos a un conjunto de procedimientos para alcanzar un determinado objetivo, el hacer uso de una

metodología no garantiza el éxito pero sí disminuye las probabilidades de fracaso.

CRISP-DM se caracteriza porque sigue el ciclo de vida de un proyecto de minería de datos, permitiendo que podamos retroceder o avanzar de una fase a otra sin complicaciones, lo contrario a lo que sucede en un modelo de cascada. Además, se parte de la premisa que el proyecto no acaba cuando se encuentra un modelo idóneo, sino que éste hay que mantenerlo, documentarlo, mejorarlo ya que puede ser usado en proyecto futuros.

En esta metodología encontramos seis fases [10], la cuales se detallan a continuación:

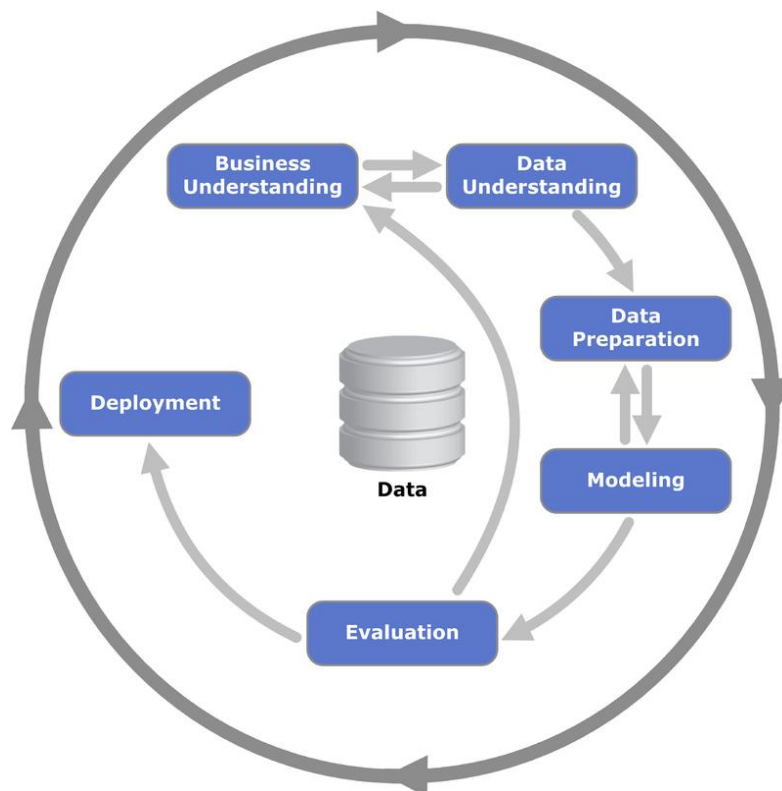


Ilustración 2 - Cross Industry Standard Process for Data Mining (CRISP-DM) [10].

1. Comprensión del negocio, en esta primera fase se busca definir los objetivos del proyecto. Para ello hay que analizar la situación en la que nos encontramos, lo que se busca resolver, qué tareas se van a realizar... Todo ello ha quedado definido en el apartado 1.3 de este documento.
2. Comprensión de los datos, posterior a la comprensión del negocio tenemos que hacer frente a qué datos tenemos, cómo son los mismos... En nuestro caso, principalmente vamos a trabajar con imágenes, que éstas son las radiografías de tórax.
3. Preparación de los datos, una vez que ya tenemos las radiografías hay que prepararlas para poder introducirlas a la red neuronal, ya sean para normalizar el tamaño de las imágenes, mejorar el contraste o el brillo...

4. Modelado, en esta fase nos encargaremos de definir la red neuronal convolucional, la cual tiene que ser capaz de resolver el problema para así poder alcanzar nuestro objetivo, en nuestro caso, diagnosticar si el paciente padece neumonía.
5. Evaluación del modelo, en ella se busca saber cómo de bueno es nuestro modelo a partir de diferentes medidas. Si la calidad del modelo no es elevada, hay que retroceder a pasos anteriores.
6. Despliegue, en esta fase el objetivo es poder transmitir de forma adecuada la conocimiento obtenido, es decir, explicar de forma comprensible cómo de bueno es el modelo generado para diagnosticar una neumonía.

1.5 Planificación del Trabajo

En el diagrama de *Gantt* que se presenta en la siguiente página, podemos apreciar el cómo se han definido las diferentes fases del proyecto junto con las fechas límites para entregar cada práctica de evaluación continua.

Cabe destacar que cada fase tiene un color diferente, el color azul representa las tareas definidas en la fase “Definición del proyecto”, el color verde indica las tareas de la fase “Estado del arte”, el color morado hace referencia a la fase “Diseño e implementación del trabajo”, el color naranja nos indica “Redacción de la memoria” y el color granate es la fase de “Defensa”.

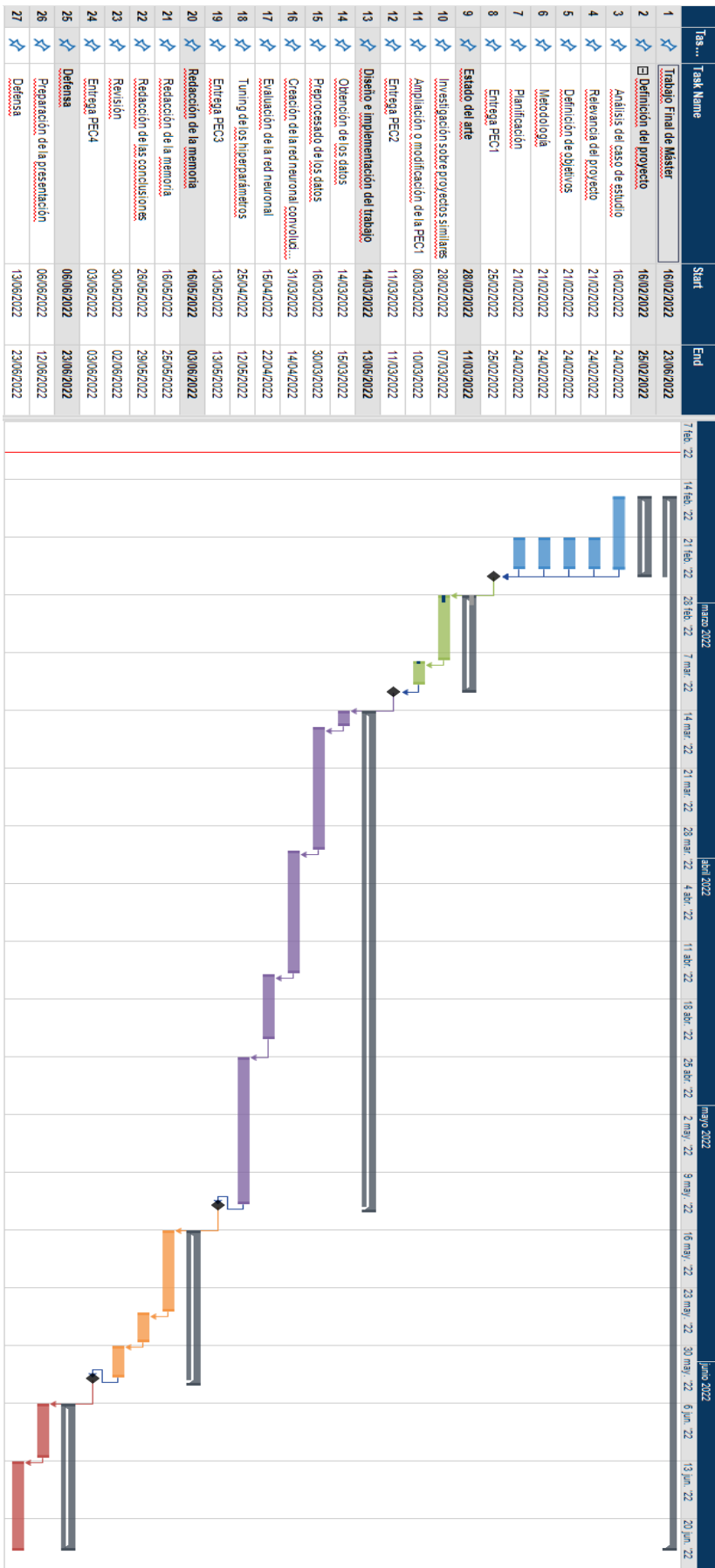


Ilustración 3 - Planificación del proyecto.

1.6 Breve resumen de productos obtenidos

A completar a medida que avanza el proyecto, añadir los entregables de las PEC junto con la presentación de la defensa.

1.7 Breve descripción de los otros capítulos de la memoria

A completar a medida que avanza el proyecto.

2. Estado del arte

2.1 Marco teórico sobre la neumonía

Gracias a los diferentes avances que se han realizado en el ámbito de la salud y en el tecnológico, permiten mejorar tanto la calidad como la esperanza de vida de las personas. Aunque esto supone un gran cambio en la sociedad, la realidad es que seguimos teniendo muchas enfermedades en el día a día. Debido a esto se tiene que seguir mejorando las diferentes técnicas para poder detectar patologías y poder tratar las mismas de una forma adecuada, con el fin de mejorar tanto la esperanza como la calidad de vida.

Una de las enfermedades más comunes es la neumonía, ésta presenta una alta incidencia tanto en los países desarrollados como en los que no lo están, dejando a su paso miles de víctimas año tras año. Además de la mortalidad que presenta la misma, ésta genera un gran impacto en la sociedad debido a la frecuencia con la que es diagnosticada, y el coste económico que supone detectar y tratar esta patología.

Por otro lado, hacer una correcta clasificación de esta enfermedad es de gran dificultad debido a que hay otras enfermedades pulmonares que presentan un cuadro similar, pero gracias a la tecnología y al personal médico se consigue identificar y tratar la misma de forma eficaz.

La neumonía es una lesión inflamatoria pulmonar como consecuencia de la llegada de microorganismos a la vía aérea distal y parénquima [11]. Para hacer un buen diagnóstico y tratamiento sobre ella, hay que tener presente el cómo evoluciona y cuáles son los síntomas que ésta presenta.

Aunque hay enfermedades pulmonares con un cuadro clínico similar, la neumonía se caracteriza porque suele presentar todos o algunos de los siguientes síntomas: tos, expectoración purulenta o herrumbrosa, disnea, dolor pleurítico y fiebre [11].

Viendo los síntomas anteriores vemos que éstos no son solo característicos de la neumonía, es más, según un estudio por tres médicos que desconocían el diagnóstico de 52 pacientes (24 con neumonía), manifestó que entre el 47% y el 69% de los casos con dicha patología fueron diagnosticados [11]. Como podemos apreciar el porcentaje de sensibilidad no es muy alto, y es aquí donde entra en juego el objetivo de este proyecto, hacer que a través del *deep learning* podamos diagnosticar más fácilmente esta enfermedad.

Por otro lado, tenemos que hablar de las diferentes neumonías que existen, y cómo estas se clasifican dependiendo de diferentes factores.

Dependiendo del agente causal, entendiendo por el mismo cualquier sustancia viva o no, cuya presencia o ausencia determina la causa de una patología [12], presentamos la siguiente clasificación [11] [13]:

- Neumonía neumocócica: se caracteriza porque la neumonía tiene origen a partir de una bacteria denominada *Streptococo pneumoniae*, conocida comúnmente como neumococo [14].
- Neumonía estafilocócica: tiene como origen una bacteria denominada *Staphylococcus aureus*, o mayoritariamente conocido como estafilococo [15]. Es relativamente raro padecer este tipo de neumonía, aunque en época de epidemia suele aumentar la importancia.
- Neumonía por *Klebsiella*: se produce a partir de la bacteria *Klebsiella*, la cual tiene consecuencias graves sobre todo entre pacientes diabéticos y alcohólicos [16].
- Neumonía por *Legionella*: es originada también por una bacteria, en este caso la bacteria *Legionella pneumophila*, un elevado porcentaje de casos con este tipo de neumonía necesitan ser tratados en unidades de cuidados intensivos, dando lugar a una elevada mortalidad [17].

Otro tipo de clasificación es dependiendo de la afectación anatomopatológica, en este caso tenemos los siguientes tipos de neumonía [11] [13]:

- Neumonía alveolar: la zona afectada mayoritariamente son los alvéolos ya que la zona de los bronquiolos no suele presentar grandes daños. Esta neumonía es la manifestación de la neumonía neumocócica.
- Neumonía multifocal: afecta tanto a los alvéolos como a los bronquiolos, pero no suele afectar a todo un lóbulo como sucedía con la anterior neumonía. Este tipo de neumonía es el claro ejemplo de neumonía estafilocócica.
- Neumonía intersticial: en este caso los alvéolos y bronquiolos no se ven afectados, sino que la zona dañada es el intersticio.
- Neumonía necrotizante: este tipo de neumonía es una complicación grave debido a que produce necrosis, muerte de un conjunto de células en el tejido pulmonar, haciendo que la mortalidad en los niños que padecen este tipo de neumonía sea elevada. La neumonía necrotizante está asociada tanto a neumonías neumocócicas como estafilocócicas.

En cuanto a la reacción de la persona que padece de dicha enfermedad, se presenta la siguiente clasificación [13]:

- Neumonía supurada: aquella neumonía que tras su recuperación no deja secuelas.

- Neumonía fibrinosa: neumonía en la que no se recupera el paciente al cien por cien, ya que tras su “recuperación” le queda un exceso en el tejido pulmonar, o también conocido como fibrosis.

Dependiendo del tipo de huésped, la neumonía puede ser [11] [13]:

- Neumonía en paciente inmunocompetente: si es capaz de dar una respuesta inmunitaria de una forma normal.
- Neumonía en paciente inmunodeprimido: si no es capaz de dar una respuesta inmunitaria de forma normal, sino que necesita de ayuda médica para hacer frente a una enfermedad.

Por último, dependiendo de dónde se ha adquirido la bacteria que origina una neumonía [11] [13]:

- Neumonías extrahospitalarias: si la patología proviene de un microorganismo ajeno al ámbito hospitalario.
- Neumonías hospitalarias: si la neumonía proviene a partir de un microorganismo del ámbito hospitalario.

2.2 Marco teórico sobre las redes neuronales convolucionales

La red neuronal convolucional (RNC) no es más que un tipo de red neuronal artificial, en otras palabras, es una técnica de *deep learning* [18]. Se caracteriza porque la RNC no necesita que se le introduzca datos, tal y como sucede en otras redes para solventar problemas de clasificación o regresión, sino que ésta es capaz de extraer características por sí sola.

Al conseguir extraer características por sí misma, hace que este tipo de red sea muy usada para encontrar patrones en imágenes, es decir, reconocer caras, personas, animales, coches, matrículas, números... Por lo tanto, el uso de este tipo de red neuronal ha aumentado en los últimos años en el ámbito de la inteligencia artificial, haciendo un mayor uso en los campos de la medicina y en la visión artificial.

Cabe destacar que aunque el tratamiento de imágenes es la principal utilidad que se le da a las RNC, éstas también son capaces de clasificar datos de audio, series temporales...

En resumen, las RNC destacan básicamente por dos motivos: el primero es que consiguen extraer características por ellas mismas, y el segundo porque son capaces de reconocer patrones con una alta precisión.

Al igual que sucede con una red neuronal artificial común, las RNC también presentan una capa de entrada y una capa de salida, es decir, el *input* de la red va a ser una imagen y el *output* la clasificación del misma. Sin embargo, en lo que difieren está en las capas intermedias entre la entrada y la salida de la red.

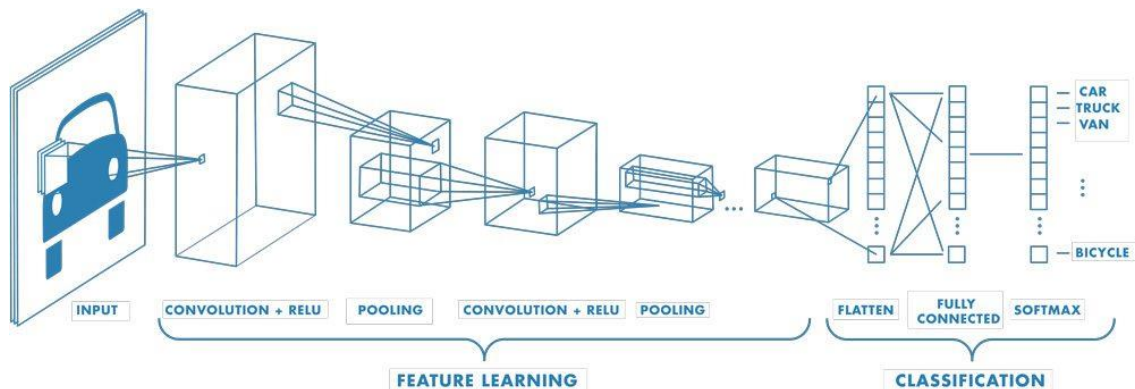


Ilustración 4 - Marco teórico sobre la red neuronal convolutiva (RNC).

Como podemos apreciar en la anterior ilustración, una red neuronal convolutiva la podemos dividir en la parte de aprendizaje de características y en la de clasificación.

Dentro de la parte de aprendizaje de características se encuentran las siguientes capas:

- Capa de convolución: es la encargada de extraer características a partir de la imagen de entrada, gracias a los filtros convolucionales.
- Capa unidad lineal rectificadora (ReLU) o función de activación: esta capa permite que solamente las características activas pasen a la siguiente fase, ya que ReLU transforma los valores negativos a cero y los valores positivos toma el valor máximo entre cero y el valor correspondiente.
- Pooling: simplifica la fase de convolución generalizando los valores, para ello por norma general se coge de un conjunto de píxeles el valor máximo.

Una característica importante de las RNC, es que las capas mencionadas anteriormente se pueden repetir tantas veces como queramos. Suele servir de gran ayuda para mejorar la precisión a la hora de clasificar.

La otra parte de la RNC es la red en sí, es decir, la que nos permite la clasificación. Dentro de ésta tenemos las siguientes capas:

- Flatening: es la capa encargada de transformar la matriz que genera la capa anterior (pooling), a un vector que sirva como entrada para la red neuronal.
- Red neuronal (fully-connected): se trata de la propia red en sí, como si fuera una red neuronal artificial común, con la característica de que todas las neuronas de una capa tienen que estar conectadas con todas las neuronas de la siguiente capa.

- La función de activación: se puede utilizar tanto la función *softmax*, la función sigmoide... Ésta nos permite dar la salida de la red neuronal y completar así la tarea de clasificación.

2.3 Estudios similares

El ámbito de la medicina siempre ha sido objeto de estudio, principalmente para mejorar la calidad y la esperanza de vida de las personas. Esto se ha conseguido gracias a los avances tanto en investigaciones medicinales como en la mejoría de la tecnología usada.

En la última década ha aparecido un nuevo factor, el *deep learning*, haciendo que los ordenadores sean capaces de aprender, hasta tal punto que nos permiten determinar si una persona padece una enfermedad, cómo prevenir la misma, qué línea de tratamiento seguir...

Por lo tanto, debido al auge que hay en la actualidad respecto al ámbito medicinal y tecnológico, ha hecho que haya numerosas investigaciones publicadas, en cuanto a clasificación de enfermedades a partir de imágenes con redes neuronales artificiales. En esta sección vamos a analizar algunas de las líneas de investigación en esta área.

Para visualizar cómo de importante está siendo este campo en los últimos años podemos hacer una búsqueda en *PubMed* [19], la cual es una base de datos que contiene millones de referencias bibliográficas sobre investigaciones en la ciencia de la salud, con la consulta "*classification deep convolutional neural networks*" podemos ver el aumento que ha habido en los últimos años, sobre todo en el año 2020 bien conocido como año del coronavirus.

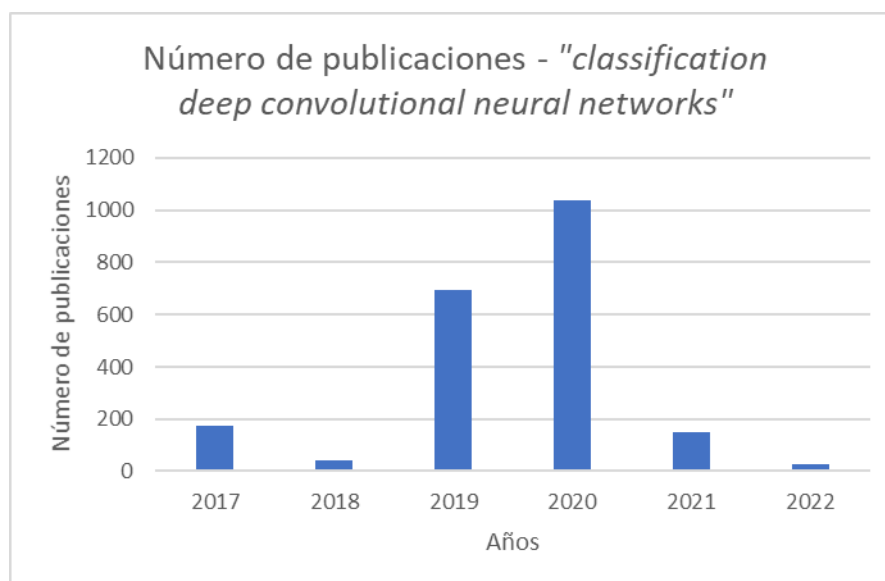


Ilustración 5 - Número de publicaciones en PubMed con la consulta "*classification deep convolutional neural networks*".

Un ejemplo del estado del arte en esta área sería el de *Samir S. Yadav* y colaboradores [20] publicado en 2019. En esta investigación se busca clasificar

imágenes médicas a partir de redes neuronales artificiales y de clasificadores de máquinas de soporte, principalmente se hace esto porque se considera que se ha alcanzado el techo con los métodos tradicionales en cuanto a rendimiento. Para poder resolver esta problemática se hace uso de dos redes neuronales convolucionales y de clasificadores de máquinas de soporte, y el objetivo es clasificar si un paciente presenta neumonía o no a partir de radiografías de tórax. El resultado de este estudio muestra que el aumento de datos mejora el rendimiento a la hora de clasificar en cualquier método, pero por otro lado, ofrece una mayor precisión una red neuronal convolucional que un clasificador de máquina de soporte, ya que se alcanza una precisión del 89% y del 77% respectivamente.

Otra de las investigaciones publicadas corresponde a *Rahib H. Abiyev* y colaboradores [21] publicado en 2021. El objetivo de esta investigación es poder diagnosticar si una persona padece SARS-Covid2 y neumonía a partir de radiografías de tórax. Para resolver esta situación se hace uso de redes neuronales convolucionales, para ser más exactos se hace uso de dos redes neuronales convolucionales entrenadas con diferentes conjuntos de datos. Uno de los modelos fue entrenado con una clasificación binaria (si el paciente presenta COVID-19 o no), mientras que el otro modelo se “nutría” del conocimiento del primer modelo (transferencia del aprendizaje). El resultado obtenido es que el modelo final consigue una precisión del 98%, haciendo que a partir de una radiografía de tórax se pueda clasificar casi con certeza si un paciente presenta neumonía y por lo tanto tiene el virus SARS-Covid2.

Si nos centramos en la clasificación de neumonías a partir de radiografías de tórax, pero sin hacer uso de redes neuronales convolucionales, tenemos la línea de investigación de *Andrés Felipe Romero* y colaboradores [22] publicada en 2020. En este estudio se realiza una segmentación de la imagen de forma manual para extraer características, posteriormente éstas se introducen a un clasificador *k-means* permitiendo así determinar si un paciente padece neumonía o no. Los resultados obtenidos son una especificidad del 76% y una sensibilidad del 83%, se consiguen unos resultados decentes sin necesidad de una red neuronal convolucional.

Otra estudio publicado sobre la clasificación de neumonías es el de *Hicham Moujahid* y colaboradores [23] publicado en 2020. En este estudio el objetivo sigue siendo el mismo, pero para resolver la problemática se hace uso de redes neuronales convolucionales. Cabe destacar que para mejorar la precisión utiliza la técnica de transferencia de aprendizaje de una red neuronal a otra, tal y como se ha indicado en estudios previos. Gracias a las redes neuronales convolucionales y a la técnica de transferencia de aprendizaje se consigue una precisión del 96%.

El último estudio que se presenta es de *Lingzhi Kong* y colaboradores [24] publicado en 2021. En este caso se busca clasificar si un paciente infantil presenta indicios de neumonía. Para ello se hace uso de una red neuronal convolucional profunda, utilizando *Xception* y una memoria tanto a corto plazo como a largo plazo (*LSTM*). *Xception* es el primer modelo, y éste se encarga de extraer las características de los datos, posteriormente *LSTM* es alimentado a

partir de la salida de *Xception*, haciendo que *LSTM* seleccione las características más importantes. La precisión alcanzada con este modelo es del 96%.

En resumen, como podemos apreciar hay numerosas líneas de investigación dentro de la ciencia de la salud, ya sean para diagnosticar cáncer, enfermedades pulmonares, genéticas... Viendo la gran cantidad de estudios que se están realizando, nos podemos hacer una idea de cómo de importante está siendo esta área. Debido a esto, en los siguientes capítulos vamos a hablar de nuestra solución a esta problemática y así aportar nuestro granito de arena.

3. Desarrollo

Tal y como ya se ha mencionado anteriormente, la metodología usada para el desarrollo de este proyecto fue *Cross Industry Standard Process for Data Mining (CRISP-DM)* [9].

Por lo tanto, las fases que definimos a continuación siguen las pautas definidas por dicha metodología, es decir, el dónde se obtienen los datos, entender los mismos, prepararlos y transformarlos para un uso posterior, crear los diferentes modelos, evaluar dichos modelos, y finalmente, preparar el modelo definitivo para poder utilizarlo en un futuro.

3.1 Extracción y compresión de los datos

3.1.1 Dónde y cómo se obtuvieron los datos

Todos los datos utilizados en este proyecto han sido obtenidos de un repositorio de *Kaggle*, para ser más exactos del repositorio “*RSNA Pneumonia Detection Challenge*” [7].



Ilustración 6 - "RSNA Pneumonia Detection Challenge" repositorio de Kaggle.

Como podemos apreciar en la anterior ilustración, éste no es un repositorio cualquiera sino que fue una competición creada por *Radiological Society of North America* [25] hace cuatro años.

El objetivo de esta competición era crear un algoritmo capaz de detectar si un paciente presentaba una neumonía a partir de imágenes médicas. Tal y como se ha comentado en la introducción de este proyecto, las neumonías presentan más del 15% de las muertes en niños menores de 5 años a nivel mundial, en el 2015 en Estados Unidos esta enfermedad causo más de 50000 muertes, dando lugar a la neumonía como una de las diez principales causas de muerte en los Estados Unidos [7].

Aunque diagnosticar con precisión la neumonía no es una tarea fácil, principalmente porque se requiere de la revisión de un especialista, el primer paso para detectar dicha enfermedad es hacer una radiografía de tórax. Las radiografías de tórax nos permiten saber si hay neumonía a partir de opacidades en los pulmones. Sin embargo, esto no siempre significa que una opacidad sea una neumonía, ya que dicha opacidad puede ser causada a partir de otras afecciones pulmonares, como por ejemplo: cáncer de pulmón, sobrecarga de líquidos, sangrado... Por lo tanto, saber si un paciente padece una neumonía con solo una radiografía es muy complicado, se necesita más información como los síntomas clínicos y un historial para afirmar el correcto diagnóstico.

Por lo tanto, *Radiological Society of North America* [25] lanzó este *challenge* junto a *US National Institutes of Health* [26], *The Society of Thoracic Radiology* [27] y *MD.ai* [28] para desarrollar un buen *dataset* de radiografías de tórax y así desafiar a la comunidad de *machine learning* en *Kaggle*.

Para la obtención de los datos hay que dirigirse al repositorio original en *Kaggle*, “*RSNA Pneumonia Detection Challenge*” [7]. Una vez dentro, tenemos que irnos a la pestaña “*Data*”, y para descargarlos tenemos tres opciones:

- Hacer uso de la *API* de *Kaggle*. En este caso, hay que introducir el siguiente comando en la consola: “*kaggle competitions download -c rsna-pneumonia-detection-challenge*”.

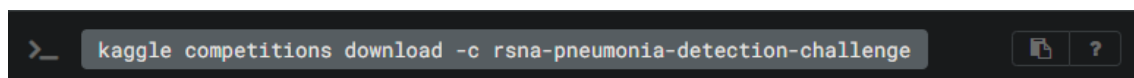



Ilustración 7 - Comando de descarga del repositorio.

- La segunda opción es descargar los archivos de uno a uno de forma manual, para ello hay que seleccionar el archivo a descargar y finalmente pulsar sobre el botón .

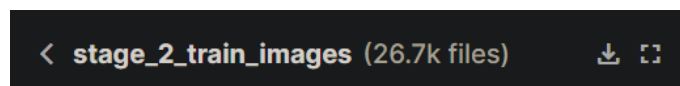


Ilustración 8 - Descarga manual de los ficheros del repositorio (uno a uno).

- La última opción es descargar los ficheros de forma manual, pero en este caso todos a la vez. Para ello hay que pulsar el botón “*Download All*”:

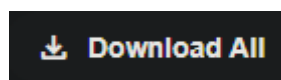


Ilustración 9 - Descarga manual de los ficheros del repositorio (todos a la vez).

Cabe mencionar que para descargar los archivos hay que cumplir dos condiciones, la primera es que debemos de tener una cuenta en *Kaggle* (no tiene ningún coste monetario asociado), y la segunda es firmar un “contrato”. Básicamente lo que se busca con dicho contrato es que se va a hacer un uso

correcto de los datos, asegurando la confidencialidad, que no se va a vender a terceros...

Una vez hecho el paso anterior podremos proceder con la descarga de los archivos y continuar con el proyecto que se desarrolla en el presente documento.

3.1.2 Estructura del repositorio

En el repositorio oficial de *Kaggle* [7] encontramos el siguiente *dashboard*:

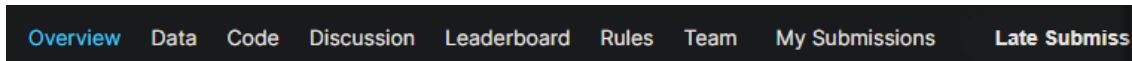


Ilustración 10 - *Dashboard* repositorio de *Kaggle*.

De la anterior imagen nos vamos a centrar en tres puntos:

- *Overview*, en él se plantea la problemática que se quiere resolver con este desafío.
- *Data*, en esta pestaña accedemos a los datos que se usan para el desarrollo de este proyecto.
- *Code*, soluciones de otros usuarios de *Kaggle* para la misma problemática o similares.

Dentro de la pestaña “*Data*” accedemos a los datos, y en ella presentamos la siguiente estructura:



Ilustración 11 - Estructura del repositorio de *Kaggle*.

- “*stage_2_test_images*”, carpeta en la que se encuentran las imágenes para test. Cabe destacar que de estas imágenes se desconoce la variable objetivo, es decir, no sabemos para ninguna de ellas si dicho paciente padece neumonía o no, es por ello que estas imágenes no han sido usadas para evaluar el modelo con el conjunto de test.
- “*stage_2_train_images*”, en ella se encuentran las imágenes para *train*. En este caso sí que conocemos la variable objetivo para cada una de las imágenes, debido a esto usaremos estas imágenes para definir los conjuntos de entrenamiento, validación y test.
- “*GCP Credits Request Link – RSNA.txt*”, es un fichero de texto que incluye una dirección web para solicitar los créditos de GCP.

- “*stage_2_detailed_class_info.csv*”, archivo CSV que recoge información sobre cada uno de los pacientes de forma detallada.
- “*stage_2_sample_submission.csv*”, archivo CSV que muestra el cómo debe subirse el resultado de la predicción para la competición. En nuestro caso, este fichero es innecesario ya que la competición está finalizada.
- “*stage_2_train_labels.csv*”, archivo CSV que contiene la información sobre los pacientes para el entrenamiento del modelo. Este fichero es fundamental para obtener la variable objetivo de cada paciente, es decir, de cada imagen.

Por otro lado, se encuentra el repositorio de *GitHub* en el que se ha desarrollado el proyecto [29]. En él encontramos la siguiente estructura:

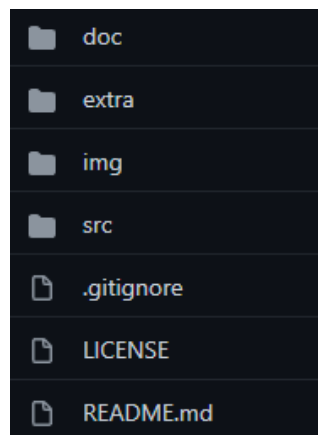


Ilustración 12 - Estructura del repositorio de *GitHub*.

- “*doc*”, carpeta en la que se encuentra la documentación para este proyecto.
- “*extra*”, carpeta en la que se encuentran ficheros usados en el desarrollo de la memoria, como la planificación de la misma...
- “*img*”, carpeta usada para almacenar las imágenes incluidas en esta memoria.
- “*src*”, carpeta en la que se encuentran los *notebooks* usados para el desarrollo de este proyecto. Cabe destacar que hay dos tipos de *notebooks*: “*Preprocessing.ipynb*” (donde se ha realizado todo el preprocesado de los datos de forma local), “*pneumonia_detection_BX*” (cada uno de los *notebooks* para la creación, entrenamiento y evaluación de cada modelo. Destacar que hay un *notebook* por cada una de la redes preentrenadas de *EfficientNet*, en este caso ejecutándose en *Google Colab* [30]).

- “*.gitignore*”, archivo en el que indicamos a *GitHub* que no suba todo al repositorio, como por ejemplo archivos temporales. Si se abre este fichero encontramos que se ignora la subida de la carpeta “*data*”, en ella se encuentran todos los archivos que se han mencionado del repositorio de *Kaggle*. Se tomó esta decisión ya que las imágenes ocupaban muchos *gigabytes*, y éstas se pueden obtener como bien se mencionó en el punto anterior.
- “*LICENSE*”, archivo en el que se indica la licencia del proyecto, en nuestro caso “*GPL-3.0 license*”. Se usó esta licencia ya que permite a otros usuarios hacer uso del proyecto sin ningún problema favoreciendo a la comunidad de *machine learning*.
- “*README.md*”, archivo de *markdown* en el que definimos la portada de nuestro repositorio.

3.1.3 Estructura de los datos

En este punto se va a comentar la información relativa a los datos en sí del proyecto.

Antes de comenzar a hablar de los ficheros, hay que mencionar dos aspectos importantes sobre las imágenes. La primera es que las imágenes no están en formato “*.PNG*” sino que están en “*DICOM*”.

El formato “*DICOM*” es el estándar que se usa para imágenes médicas ya que es un protocolo estándar de comunicación y almacenamiento de imágenes [31]. La clave de este formato es que aparte de la imagen en sí, se puede almacenar información relativa al paciente, como por ejemplo: el sexo, estatura, peso... Esta información adicional no se ha hecho uso de ella ya que no es necesaria para la resolución del problema.

El otro punto importante sobre las imágenes es que en el formato original cada imagen tiene un tamaño de 1024x1024 píxeles, el cual se tiene que transformar para hacer uso de las redes neuronales que definiremos más adelante.

Respecto a los ficheros, tal y como hemos mencionado antes hay dos, “*stage_2_detailed_class_info.csv*” y “*stage_2_train_labels.csv*”.

En el primero de ellos encontramos la siguiente estructura:

patientId	class
26684 unique values	No Lung Opacity / ... 39% Lung Opacity 32% Other (8851) 29%
0004cfab-14fd-4e49-80ba-63a80b6bddd6	No Lung Opacity / Not Normal
00313ee0-9eaa-42f4-b0ab-c148ed3241cd	No Lung Opacity / Not Normal

Ilustración 13 - Estructura de "stage_2_detailed_class_info.csv".

- "patientId": este atributo hace referencia al id del paciente, dentro de la carpeta de imágenes si buscamos dicho atributo más el formato *DICOM* ".DCM" encontramos la imagen correspondiente.
- "class": en este caso el atributo hace referencia a si el paciente pertenece una de las siguientes tres clases: *No Lung Opacity / Not normal*, *Normal*, *Lung Opacity*. La dos primeras clases nos indican que no hay neumonía en la imagen, mientras que la tercera indica lo contrario.

Para resolver el problema que se plantea en este proyecto se podría haber hecho uso del anterior fichero, pero en nuestro caso no ha sido así. Se decidió hacer uso de las clases definidas en el fichero "stage_2_train_labels.csv" básicamente porque nos interesaba saber si hay neumonía o no, es decir, si es un 1 o un 0. Dicho fichero presenta los siguientes atributos:





patientId	# x	# y	# width	# height	# Target
26684 unique values					
0004cfab-14fd-4e49-80ba-63a80b6bddd6					0
00313ee0-9eaa-42f4-b0ab-c148ed3241cd					0
00322d4d-1c29-4943-afc9-b6754be640eb					0
003d8fa0-6bf1-40ed-b54c-ac657f8495c5					0
00436515-870c-4b36-a041-de91049b9ab4	264.0	152.0	213.0	379.0	1
00436515-870c-4b36-a041-de91049b9ab4	562.0	152.0	256.0	453.0	1

Ilustración 14 - Estructura de "stage_2_train_labels.csv".

- "patientId", al igual que antes indica el id del paciente, necesario para localizar la imagen.
- "x", tanto este atributo como los siguientes tres son usados para identificar el área donde hay una opacidad pulmonar. Éste en concreto indica dónde empieza el área de la opacidad pulmonar respecto al eje de la x.

- “*y*”, indica dónde empieza la opacidad pulmonar respecto al eje de la *y*.
- “*width*”, anchura de la opacidad pulmonar.
- “*height*”, altura del área de la opacidad pulmonar.
- “*Target*”, variable objetivo. En nuestro caso, si hay un 1 hay opacidad pulmonar (representa a la clase *Lung Opacity*), si hay un 0 no hay neumonía (representa tanto a la clase *No Lung Opacity / Not Normal* como a la clase *Normal*).

3.2 Preprocesado de los datos

3.2.1 Lectura y filtrado de los datos

El primer paso en la preparación de los datos es la lectura de los mismos, para ello se ha hecho uso de la librería de *pandas* tanto para leer el archivo “*stage_2_train_labels.csv*” como para crear el *dataframe* correspondiente y manipularlo.

Una vez hecha la lectura del fichero, obtendremos 30227 registros y 6 dimensiones, las cuales, son los campos definidos en el punto anterior del presente documento (“*patientId*”, “*x*”, “*y*”, “*width*”, “*height*”, “*Target*”). El resultado de esta lectura sería el siguiente:

	<i>patientId</i>	<i>x</i>	<i>y</i>	<i>width</i>	<i>height</i>	<i>Target</i>
0	0004cfab-14fd-4e49-80ba-63a80b6bddd6	NaN	NaN	NaN	NaN	0
1	00313ee0-9eaa-42f4-b0ab-c148ed3241cd	NaN	NaN	NaN	NaN	0
2	00322d4d-1c29-4943-afc9-b6754be640eb	NaN	NaN	NaN	NaN	0
3	003d8fa0-6bf1-40ed-b54c-ac657f8495c5	NaN	NaN	NaN	NaN	0
4	00436515-870c-4b36-a041-de91049b9ab4	264.0	152.0	213.0	379.0	1

Ilustración 15 - Lectura de los datos.

A continuación comprobamos si hay elementos duplicados, en nuestro caso vemos que sí que los hay. Un aspecto a destacar de estos duplicados es que en sí no son duplicados, es decir, en los registros cuyo valor de “*Target*” sea 1, en otras palabras en los que hay neumonía, sí que hay elementos duplicados; básicamente porque se refleja tanto la opacidad pulmonar para el pulmón izquierdo como para el derecho.

En nuestro problema, queremos identificar los casos en los que hay neumonía independientemente dónde se encuentre la misma, es decir, si es en un pulmón o en el otro. Es por ello que eliminamos los duplicados, seleccionando el primer registro y descartando el segundo.

Como resultado de esta operación, pasamos de tener 30227 registros a 26684, es decir, se han descartado 3543.

El siguiente paso es filtrar las columnas, básicamente quedarnos con los datos que son necesarios y descartando los demás. Por lo tanto, de 6 dimensiones que tiene el *dataset* ("*patientId*", "*x*", "*y*", "*width*", "*height*", "*Target*") nos quedamos con 2 ("*patientId*" y "*Target*").

3.2.2 Creación de los conjuntos de entrenamiento, validación y test

Una vez que hemos realizado el preprocesado de los datos continuamos con la creación de los conjuntos de entrenamiento, validación y test.

De los 26684 registros que tenemos en total, se ha decidido dividir dicho conjunto en 3000 registros para test y otros 3000 para validación, y 20684 para el entrenamiento.

Se tomó la decisión de esta partición porque en la carpeta "*stage_2_test_images*" del repositorio de *Kaggle* [7], se pone a disposición de 3000 imágenes para test. Así que siguiendo la misma pauta se han definido dichos conjuntos.

3.2.3 Transformación de las imágenes

En este apartado se realiza el preprocesado de las imágenes en sí, es decir, realizamos la transformación del formato "*DICOM*" al "*PNG*" y se procede con el redimensionado de las imágenes.

Respecto al cambio de formato, se determinó que lo mejor era convertir las mismas a formato "*PNG*" por dos motivos, el primero con este formato es más fácil trabajar, y en segundo lugar viene determinado por el *data augmentation* que realizaremos en el siguiente punto, ya que es más sencillo realizar el guardado de una imagen "*PNG*" que una con el formato "*DICOM*".

Por otro lado, a la par que se transforma el formato de las imágenes realizamos el redimensionado de las mismas, esto se debe a que el tiempo de entrenamiento de la red neuronal va a ser menor si el *input* que recibe la misma ya está en el tamaño adecuado.

Aunque más adelante vamos a hablar de los diferentes modelos creados para resolver este problema, a continuación se muestra en la siguiente tabla los diferentes tamaños que se necesitan para cada red neuronal:

Red neuronal	Tamaño de las imágenes
EfficientNetB0	224x224
EfficientNetB3	300x300
EfficientNetB4	380x380
EfficientNetB5	456x456

Tabla 1 - Tamaños de las imágenes para cada red neuronal.

Por último, cabe destacar que en todas las tareas de preprocesado se comprueba que todo se realiza de forma correcta. De tal forma que cuando vayamos a usar los datos para el entrenamiento de los modelos esté todo en orden.

3.2.3 Balanceo de los datos

Continuando con el preprocesado de los datos falta comprobar el balanceo de los mismos, es decir, si para el conjunto de entrenamiento, el cual vamos a usar para entrenar la red, tiene el mismo número de casos de pacientes con neumonía que sin neumonía.

Ejecutando los comandos correspondientes vemos que no hay el mismo número de casos, para ser más exactos, de 20684 imágenes tenemos 16072 sin neumonía y 4612 con neumonía. En otras palabras, hay casi un factor de 3.5 imágenes más de pacientes sin neumonía que con neumonía, por lo que hay un claro desbalanceo de clases.

Solventar el desbalanceo de clases es fundamental para el entrenamiento de la red, ya que de lo contrario podríamos tener una red con pesos sesgados, es decir, que la red se especialice más en los casos sin neumonía que con neumonía.

La solución a este problema viene dada a partir del aumento de datos de la clase minoritaria, en nuestro caso de las imágenes con neumonía.

Para ello, se ha realizado el *data augmentation* con el módulo *ImageDataGenerator* de *Keras*, gracias a él podemos crear imágenes nuevas a partir de las originales, estableciendo parámetros como: la rotación de la imagen, giro horizontal, zoom, desplazar la misma a lo ancho o vertical...

Este aumento de datos se realiza en dos fases, en la primera se van a crear tantas imágenes aumentadas como sea la diferencia con la proporción, es decir, si la proporción es 3.5 se van a crear $3 - 1 = 2$ imágenes aumentadas por cada imagen con neumonía. De esta forma nos garantizamos que todos los casos con neumonía al menos van a tener dos imágenes aumentadas.

La segunda fase consiste en crear una imagen aumentada a partir de las imágenes originales con neumonía de forma aleatoria, hasta llegar al punto en el que haya un balanceo de clases.

Una vez alcanzado el balanceo de clases aumenta el *dataset* de entrenamiento, se pasa de 20684 imágenes a 32144. De esta forma se consigue dotar de mayor calidad el *dataset* original.

En las siguientes ilustraciones se puede apreciar las imágenes que genera el *data augmentation*:

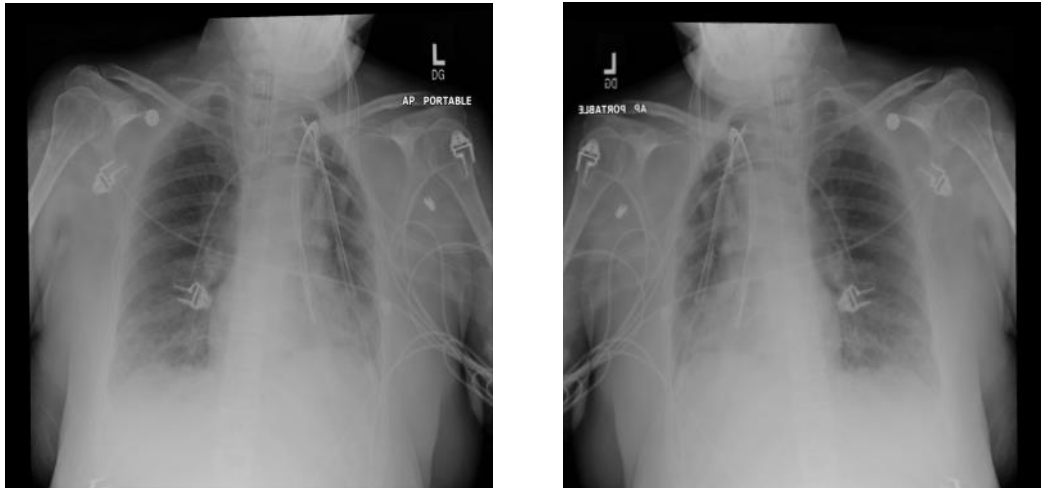


Ilustración 16 - Imagen original vs imagen aumentada.

Finalmente, cabe mencionar el nombre que se le da a las imágenes aumentadas. Para las imágenes que se crean en la fase uno su nombre viene proporcionado por el atributo “*patientId*”, más “#a”, más la iteración en la que se crea (0,1). Mientras que para la fase dos, el nombre es el atributo “*patientId*”, más “#b”, más la iteración en la que se crea (en nuestro caso 0).

En la siguiente tabla podemos apreciar tanto el nombre original como las nombres de las imágenes aumentadas:

Nombre original	<i>0a8d486f-1aa6-4fcf-b7be-4bf04fc8628b.png</i>
Fase 1 – iteración 1	<i>0a8d486f-1aa6-4fcf-b7be-4bf04fc8628b#a0.png</i>
Fase 1 – iteración 2	<i>0a8d486f-1aa6-4fcf-b7be-4bf04fc8628b#a1.png</i>
Fase 2 – iteración 1	<i>0a8d486f-1aa6-4fcf-b7be-4bf04fc8628b#b0.png</i>

Tabla 2 - Comparativa entre el nombre de la imagen original y las aumentadas.

3.2.4 Generación de los ficheros de salida

Esta es la última fase de preprocesado, en ella lo que se hace es generar todos los archivos “CSV” necesarios para continuar con el siguiente apartado de

creación de los modelos, es decir, se generan los archivos de entrenamiento, validación y test.

Estos ficheros resultantes contienen los siguientes datos:

- “*train_labels.csv*”, en él encontramos al atributo “*PatientId*” y “*Target*” sobre las imágenes de entrenamiento. Este archivo contiene 32144 registros.
- “*validation_labels.csv*”, al igual que el anterior contiene “*PatientId*” y “*Target*” para el conjunto de validación. En este caso hay 3000 registros.
- “*test_labels.csv*”, sigue teniendo los mismos atributos y en este caso hay 3000 registros.

3.3 Creación de los modelos

3.3.1 Data Generator

Una de las problemáticas de trabajar con imágenes es que a la hora de entrenar los modelos hay dificultades en cargar en memoria todas ellas. El *dataset* de entrenamiento está formado por 32144 imágenes, que dependiendo del modelo que usemos, los cuales serán explicados más adelante, dichas imágenes tendrán diferentes tamaños, haciendo que a mayor tamaño mayor consumo de memoria.

Para solventar esta situación se crea un *data generator*, el cual nos va a permitir ir cargando en memoria pequeños grupos de imágenes, o como se denomina en este campo *batch*, a medida que se vayan necesitando.

La creación de este *data generator* se consigue gracias a la clase “*Sequence*” de la librería *Keras*.

Dentro de esta clase se han definido los siguientes métodos:

- “*__init__*(...)”, este método es el constructor de la clase, el cual nos va a permitir crear cada uno de nuestros objetos de tipo “*Sequence*”.
- “*__train__*(...)”, se encarga de encontrar la imagen y cargarla en memoria, además de encontrar para cada imagen el valor de la variable objetivo. En otras palabras, nos devuelve la imagen y si ésta indica que el paciente padece neumonía o no.
- “*__test__*(...)”, en este caso se carga en memoria y se devuelve la imagen correspondiente, es decir, no hay un retorno de la variable objetivo.

- “`__getitem__(...)`”, método perteneciente a la clase “*Sequence*”, el cual a partir de un índice es capaz de generar el *batch* necesario para el entrenamiento, validación y test.
- “`__on_epoch_end()`”, este método se encarga de mezclar los nombres de los archivos para dotar de mayor aleatoriedad en el entrenamiento.
- “`__len__()`”, devuelve el número de *batches* que se tienen que generar.

3.3.2 Carga de las imágenes preprocesadas

Uno de los principales inconvenientes que se ha tenido a la hora de desarrollar este proyecto ha sido el tiempo de ejecución a la hora de entrenar los modelos.

Al principio se ideó como solución del proyecto el hacer uso exclusivamente del *data generator* para reducir el consumo de memoria, para ello dentro del propio generador se tenía que buscar la imagen, abrirla y redimensionar al tamaño correspondiente para cada modelo, esto lo que significa es mucho tiempo de ejecución, para ser más exactos, con la red más sencilla tardaba el modelo dos horas por época en entrenarse.

Para solventar esta situación ya que era inviable que tardase tanto, se decidió sacrificar el consumo de memoria, es decir, en vez de cargar las imágenes poco a poco con el *data generator* se cargan todas a la vez, y se hace uso del generador para introducir a la red cada *batch* como *input*.

Esta carga de las imágenes se realiza a partir de un diccionario. Todo diccionario está formado por una pareja clave-valor, nuestra clave es el atributo “*patientId*” (el nombre de las imágenes sin la extensión “.PNG”) y el valor es la imagen preprocesada con el tamaño correspondiente.

Aunque puede sonar contradictorio el definir el *data generator* para reducir el consumo de la memoria, y luego hacer uso de un diccionario que contiene todas las imágenes, las cuales son cargadas todas en memoria, la realidad es todo lo contrario. En los equipos que tienen una capacidad de cómputo elevada (como en el caso de *Google Colab* [30]) haciendo uso tanto del *data generator* como del diccionario se consigue reducir de forma radical el tiempo de entrenamiento, ya que donde antes tardaba dos horas por época con el modelo más sencillo ahora se tarda un minuto y medio; por otro lado, si un usuario no puede acceder a una capacidad de cómputo elevada, tiene la opción de no usar el diccionario (evitar un mayor consumo de memoria) y hacer uso solamente del *data generator*, consiguiendo poder entrenar la red, pero a cambio de un elevado tiempo de entrenamiento.

3.3.3 Creación de las redes neuronales convoluciones con transfer learning

El siguiente paso en el desarrollo del proyecto ha sido crear las diferentes redes neuronales convolucionales para resolver el problema que se ha definido en puntos anteriores.

Recapitulando el apartado 2.2 del presente documento, donde se explicaban las redes neuronales convolucionales, dichas redes se caracterizan por recibir como *input* un vector (la imagen) y proporcionar como *output* la clase a la que pertenece dicha imagen.

Crear una red neuronal convolucional desde cero es un trabajo muy costoso, tanto la creación como el entrenamiento de la misma, y es aquí donde aparece un concepto fundamental, el *transfer learning*.

El *transfer learning* [32] [33] nos va a permitir solucionar la problemática comentada en el párrafo anterior. El objetivo de este método es hacer uso de redes neuronales preentrenadas con otros conjuntos de datos, para así reducir el tiempo de entrenamiento de la red y obtener una buena precisión en los conjuntos de validación y de test.

La mecánica del *transfer learning* consiste en definir un modelo base preentrenado, y posterior a éste crear un nuevo modelo basado en el modelo base. De forma resumida lo que se busca es lo siguiente:

1. Inicializar un modelo base con pesos ya preentrenados.
2. Congelar los pesos de las capas del modelo base.
3. Creamos un nuevo modelo a partir del modelo base.
4. Entrenamos el nuevo modelo con el *dataset*.

En otras palabras, lo que se busca es entrenar el nuevo modelo, es decir el clasificador, pero haciendo uso del modelo base.

En nuestro caso se ha hecho uso de *EfficientNet* como modelo base. *EfficientNet* [34] no es más que un conjunto de redes neuronales convolucionales preentrenadas proporcionadas por *Google*. Estas redes preentrenadas presentan las siguientes capas:

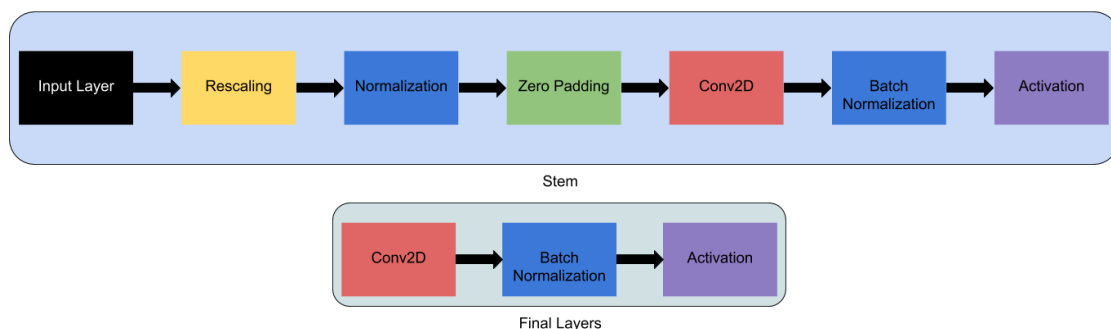


Ilustración 17 - Capas de un modelo *EfficientNet*. [35]

Todos los modelos que se han creado para resolver la problemática de este proyecto presentan la siguiente estructura:

1. Se define el *input* que va a tener la imagen, en nuestro caso depende de qué tipo de *EfficientNet* se use. Si usamos como ejemplo *EfficientNetB0* el tamaño de la imagen va a ser 224x224, y solo un canal ya que las imágenes están en escala de grises.
2. Después, para dotar de mayor variabilidad al modelo se ha creado una capa de *data augmentation*.
3. Un aspecto a tener en cuenta es que las redes *EfficientNet* hacen uso de imágenes en color, es decir, se necesitan tres canales. En nuestro caso las imágenes están en escala de grises por lo que tenemos solamente un canal. Para solventar esta situación, se ha creado una capa "*Concatenate*", concatenando la misma imagen tres veces conseguimos los tres canales.
4. Una vez realizado esto ya podemos aplicar *transfer learning*, para ello se inicializa el modelo de *EfficientNet* que queremos usar, en este proyecto se han creado modelos para *EfficientNetB0*, *EfficientNetB3*, *EfficientNetB4* y *EfficientNetB5*. Un aspecto importante en la inicialización es que las capas finales no estén incluidas en el modelo base.
5. Creado el modelo base pasamos a congelar todas las capas del mismo.
6. Posteriormente, creamos de nuevo las capas finales. Básicamente se crea una capa "*GlobalAveragePooling2D*", una capa "*BatchNormalization*", una capa "*Dropout*" y una capa de salida "*Dense*".
7. En nuestro problema, buscamos saber si un paciente padece una neumonía o no, es decir, la red devuelve una probabilidad entre 0 y 1. Es por ello, que la capa de salida "*Dense*" presenta una activación de tipo "*sigmoid*".
8. Para terminar, falta compilar el modelo. Como optimizador se ha usado "*Adam*", como función de pérdida "*binary_crossentropy*" y como métrica "*accuracy*".

En la sección 4 del presente documento, se explicará qué experimentos se han realizado, cuáles han sido los resultados y qué modelo se ha comportado mejor, pero para terminar con esta sección 3 nos falta hablar de *fine-tuning*.

3.3.4 Fine-tuning

Aplicado el *transfer learning* al modelo, tenemos que mejorar la precisión. Para ello, se hace uso de otra técnica conocida como *fine-tuning*.

Fine-tuning [33] tiene como objetivo mejorar la precisión que se obtiene al aplicar *transfer learning*, esto se consigue gracias a descongelar parte del modelo base o todo el modelo, entrenar de nuevo la red y usar un *learning rate* bajo.

Cuando se habla de “descongelar”, a lo que se refiere es a que los pesos de las capas del modelo base ahora sí que se van a modificar cuando se entrene la red, es decir, con *transfer learning* se entrena solo el clasificador final (pesos del clasificador final) y con *fine-tuning* los pesos de toda la red.

Para el desarrollo de este proyecto, a la hora de aplicar *fine-tuning* se ha decidido descongelar todo el modelo base, con la excepción de las capas “*BatchNormalization*”, ya que según establece *Keras* en su web oficial, descongelar estas capas reducen la precisión de la primera época en el *training* del modelo [36].

4. Experimentos y resultados

5. Conclusiones

Este capítulo tiene que incluir:

- Una descripción de las conclusiones del trabajo: Qué lecciones se han aprendido del trabajo?.
- Una reflexión crítica sobre el logro de los objetivos planteados inicialmente: Hemos logrado todos los objetivos? Si la respuesta es negativa, por qué motivo?
- Un análisis crítico del seguimiento de la planificación y metodología a lo largo del producto: Se ha seguido la planificación? La metodología prevista ha sido la adecuada? Ha habido que introducir cambios para garantizar el éxito del trabajo? Por qué?
- Las líneas de trabajo futuro que no se han podido explorar en este trabajo y han quedado pendientes.

6. Líneas de trabajo futuro

7. Glosario

Definición de los términos y acrónimos más relevantes utilizados dentro de la Memoria.

8. Bibliografía

- [1] «Neumonía». <https://medlineplus.gov/spanish/pneumonia.html> (accedido 22 de febrero de 2022).
- [2] «Neumonía», *Wikipedia, la enciclopedia libre*. 22 de febrero de 2022. Accedido: 22 de febrero de 2022. [En línea]. Disponible en: <https://es.wikipedia.org/w/index.php?title=Neumon%C3%ADa&oldid=141845195>
- [3] «SARS-CoV-2», *Wikipedia, la enciclopedia libre*. 16 de febrero de 2022. Accedido: 23 de febrero de 2022. [En línea]. Disponible en: <https://es.wikipedia.org/w/index.php?title=SARS-CoV-2&oldid=141723154>
- [4] «La neumonía causa 10.000 muertes anuales, muchas de ellas prevenibles», *El Independiente*, 6 de noviembre de 2019. <https://www.elindependiente.com/vida-sana/salud/2019/11/07/la-neumonia-causa-10-000-muertes-anuales-muchas-de-ellas-prevenibles/> (accedido 22 de febrero de 2022).
- [5] «Neumonía». <https://www.who.int/es/news-room/fact-sheets/detail/pneumonia> (accedido 22 de febrero de 2022).
- [6] L. Kong y J. Cheng, «Based on improved deep convolutional neural network model pneumonia image classification», *PLOS ONE*, vol. 16, n.º 11, p. e0258804, nov. 2021, doi: 10.1371/journal.pone.0258804.
- [7] «RSNA Pneumonia Detection Challenge». <https://kaggle.com/c/rsna-pneumonia-detection-challenge> (accedido 23 de febrero de 2022).
- [8] «DICOM», *Wikipedia, la enciclopedia libre*. 27 de enero de 2022. Accedido: 23 de febrero de 2022. [En línea]. Disponible en: <https://es.wikipedia.org/w/index.php?title=DICOM&oldid=141264476>
- [9] «CRISP-DM: La metodología para poner orden en los proyectos», *Sngular*, 2 de agosto de 2016. <https://www.sngular.com/es/data-science-crisp-dm-metodologia/> (accedido 24 de febrero de 2022).
- [10] J. F. V. Rueda, «CRISP-DM: una metodología para minería de datos en salud», *healthdataminer.com*, 4 de noviembre de 2019. <https://healthdataminer.com/data-mining/crisp-dm-una-metodologia-para-mineria-de-datos-en-salud/> (accedido 24 de febrero de 2022).
- [11] J. L. García Satué y J. Aspa Marco, *Neumonías*. Majadahonda (Madrid: Ergon, 2006).
- [12] «Agente (medicina)», *Wikipedia, la enciclopedia libre*. 17 de febrero de 2022. Accedido: 6 de marzo de 2022. [En línea]. Disponible en: [https://es.wikipedia.org/w/index.php?title=Agente_\(medicina\)&oldid=141730538](https://es.wikipedia.org/w/index.php?title=Agente_(medicina)&oldid=141730538)
- [13] «Neumonía», *Wikipedia, la enciclopedia libre*. 27 de febrero de 2022. Accedido: 6 de marzo de 2022. [En línea]. Disponible en: <https://es.wikipedia.org/w/index.php?title=Neumon%C3%ADa&oldid=141952610>
- [14] «Enfermedad Neumocócica (Incluye la neumonía neumocócica, la meningitis neumocócica y la bacteremia neumocócica)». https://www.health.ny.gov/es/diseases/communicable/pneumococcal/fact_sheet.htm (accedido 6 de marzo de 2022).
- [15] N. estafilocócica N. causada por un estafilococos A. transmisor: S. aureus, «Neumonía Estafilocócica - EcuRed».

- https://www.ecured.cu/Neumon%C3%ADa_Estafiloc%C3%B3cica (accedido 6 de marzo de 2022).
- [16] «Infecciones por Klebsiella, Enterobacter y Serratia - Enfermedades infecciosas», *Manual Merck versión para profesionales*. <https://www.merckmanuals.com/es-us/professional/enfermedades-infecciosas/bacilos-gramnegativos/infecciones-por-y> (accedido 6 de marzo de 2022).
- [17] J. I. Sánchez Olmedo, V. Jorge Amigo, M. Pérez Alé, J. Garnacho Montero, F. J. Jiménez, y C. Ortiz Leyba, «Neumonía comunitaria grave por Legionella pneumophila», *Med. Intensiva*, vol. 26, n.º 10, pp. 504-507, dic. 2002.
- [18] «Redes neuronales convolucionales». <https://es.mathworks.com/discovery/convolutional-neural-network-matlab.html> (accedido 6 de marzo de 2022).
- [19] «PubMed», *PubMed*. <https://pubmed.ncbi.nlm.nih.gov/> (accedido 7 de marzo de 2022).
- [20] S. S. Yadav y S. M. Jadhav, «Deep convolutional neural network based medical image classification for disease diagnosis», *J. Big Data*, vol. 6, n.º 1, p. 113, dic. 2019, doi: 10.1186/s40537-019-0276-2.
- [21] R. H. Abiyev y A. Ismail, «COVID-19 and Pneumonia Diagnosis in X-Ray Images Using Convolutional Neural Networks», *Math. Probl. Eng.*, vol. 2021, p. e3281135, nov. 2021, doi: 10.1155/2021/3281135.
- [22] A. F. R. Gómez, C. A. C. Guarnizo, A. M. G. Pomarico, y S. L. C. Suárez, «ALGORITMO PARA LA CARACTERIZACIÓN DE NEUMONÍA EN RADIOGRAFÍA DE TÓRAX», *Encuentro Int. Educ. En Ing.*, ago. 2020, Accedido: 3 de marzo de 2022. [En línea]. Disponible en: <https://acofipapers.org/index.php/eiei/article/view/794>
- [23] «Convolutional Neural Network Based Classification of Patients with Pneumonia using X-ray Lung Images», *Journal*. <https://astesj.com/v05/i05/p22/> (accedido 3 de marzo de 2022).
- [24] L. Kong y J. Cheng, «Based on improved deep convolutional neural network model pneumonia image classification», *PLOS ONE*, vol. 16, n.º 11, p. e0258804, nov. 2021, doi: 10.1371/journal.pone.0258804.
- [25] «Radiological Society of North America». <https://www.rsna.org/> (accedido 13 de mayo de 2022).
- [26] «National Institutes of Health (NIH)», *National Institutes of Health (NIH)*. <https://www.nih.gov/> (accedido 13 de mayo de 2022).
- [27] «Society of Thoracic Radiology». <https://thoracicrad.org/> (accedido 13 de mayo de 2022).
- [28] «MD.ai». <https://www.md.ai/> (accedido 13 de mayo de 2022).
- [29] M. Ubierna San Mamés, *Pneumonia-Detection*. 2022. Accedido: 13 de mayo de 2022. [En línea]. Disponible en: <https://github.com/mariouism/Pneumonia-Detection>
- [30] «Google Colaboratory». <https://colab.research.google.com/?hl=es> (accedido 13 de mayo de 2022).
- [31] «¿Qué es el formato DICOM? Las claves del estándar en imágenes médicas», *Clinic Cloud*, 22 de julio de 2021. <https://clinic-cloud.com/blog/formato-dicom-que-es-estandar-imagenes-medicas/> (accedido 13 de mayo de 2022).

- [32] K. Team, «Keras documentation: Transfer learning & fine-tuning». https://keras.io/guides/transfer_learning/ (accedido 13 de mayo de 2022).
- [33] R. Alberto, «Redes Neuronales: ¿Qué es Transfer Learning y Fine Tuning?», *Medium*, 4 de julio de 2021. <https://rubialesalberto.medium.com/redes-neuronales-qu%C3%A9-es-transfer-learning-y-fine-tuning-8259a81cfdbc> (accedido 13 de mayo de 2022).
- [34] M. Tan y Q. V. Le, «EfficientNet: Rethinking Model Scaling for Convolutional Neural Networks». arXiv, 11 de septiembre de 2020. Accedido: 13 de mayo de 2022. [En línea]. Disponible en: <http://arxiv.org/abs/1905.11946>
- [35] «Social Network for Programmers and Developers». <https://morioh.com> (accedido 13 de mayo de 2022).
- [36] K. Team, «Keras documentation: Image classification via fine-tuning with EfficientNet». https://keras.io/examples/vision/image_classification_efficientnet_fine_tuning/#tips-for-fine-tuning-efficientnet (accedido 13 de mayo de 2022).

9. Anexos

Listado de apartados que son demasiado extensos para incluir dentro de la memoria y tienen un carácter autocontenido (por ejemplo, manuales de usuario, manuales de instalación, etc.)

Dependiente del tipo de trabajo, es posible que no haya que añadir ningún anexo.