

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>
<p>Resumen del Trabajo (máximo 250 palabras): <i>Con la finalidad, contexto de aplicación, metodología, resultados i conclusiones del trabajo.</i></p> <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 sumario 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. Conclusiones.....	15
4. Glosario	16
5. Bibliografía	17
6. Anexos	19

Lista de figuras

Ilustración 1 - (a) radiografía paciente normal, (b) radiografía paciente con neumonía [6].	2
Ilustración 2 - Cross Industry Standard Process for Data Mining (CRISP-DM) [10].	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 " <i>classification deep convolutional neural networks</i> ".	12

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:

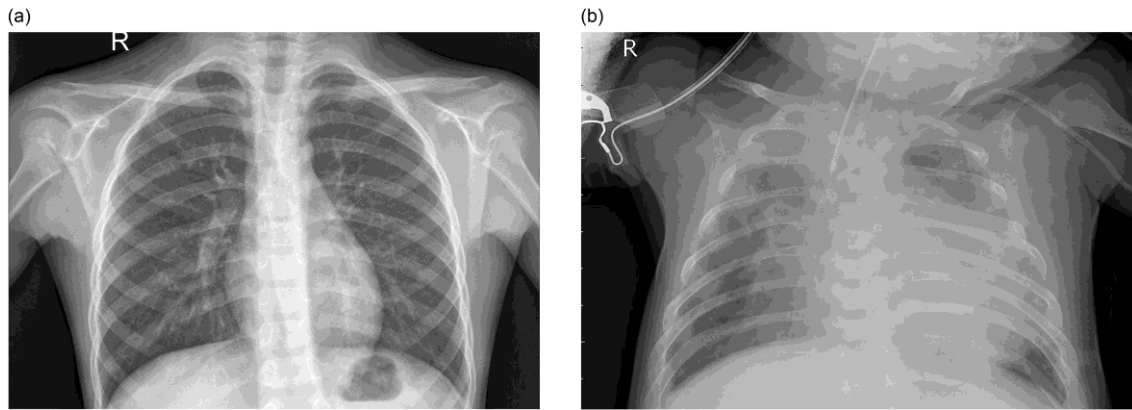


Ilustración 1 - (a) radiografía paciente normal, (b) radiografía paciente con neumonía [6].

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:

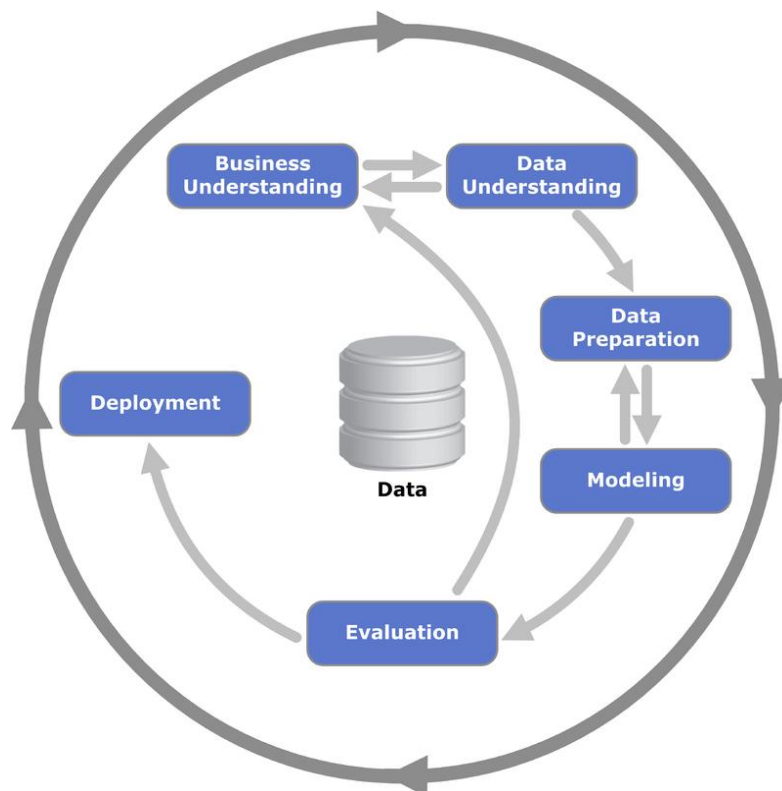


Ilustración 2 - Cross Industry Standard Process for Data Mining (CRISP-DM) [10].

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”.

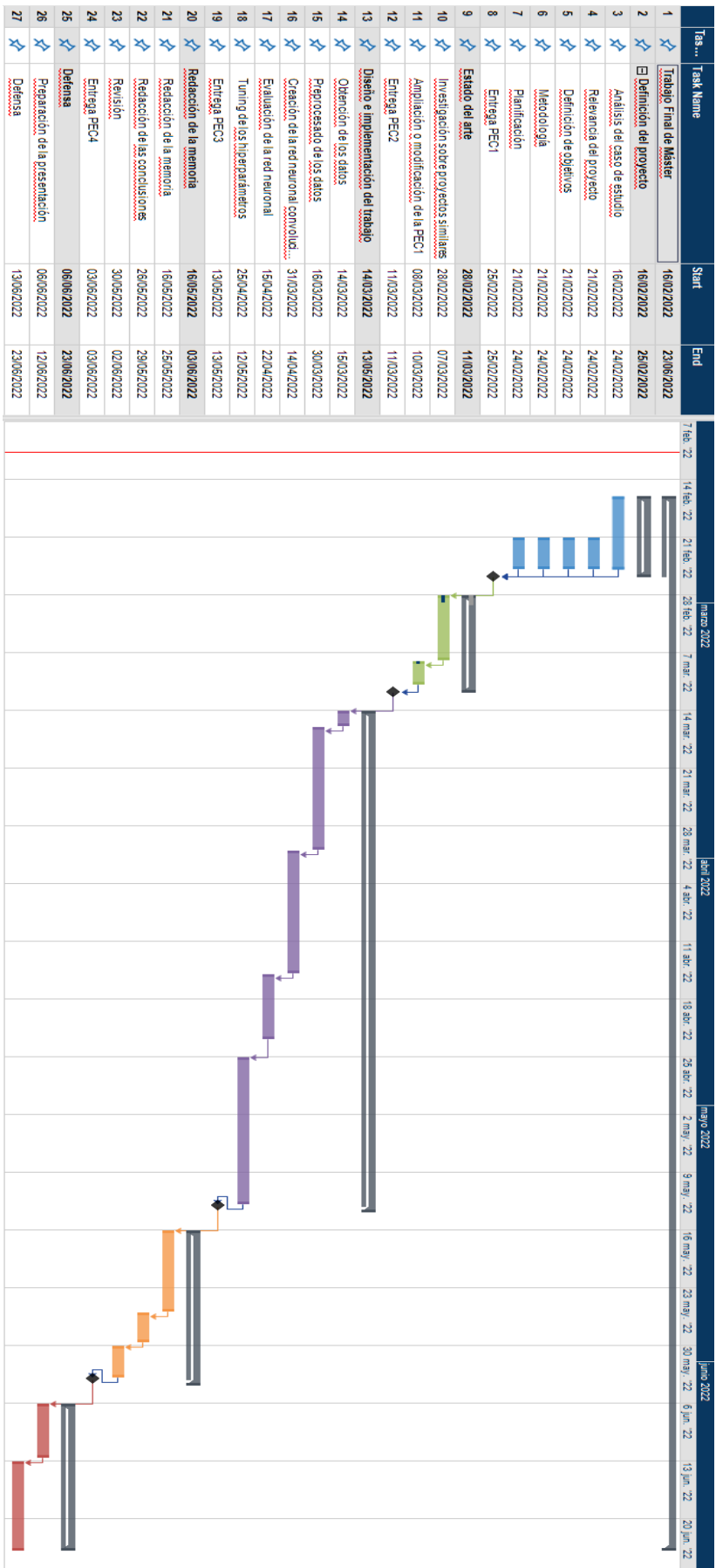


Ilustración 3 - Planificación del proyecto.

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.

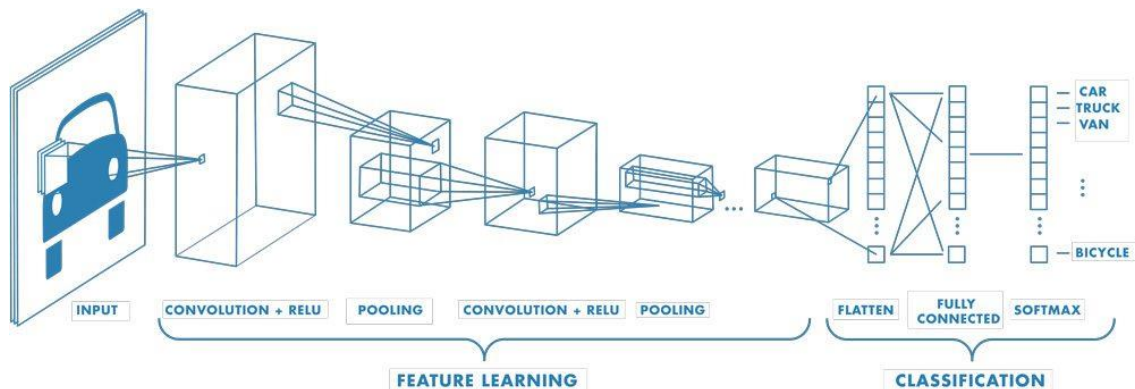


Ilustración 4 - Marco teórico sobre la red neuronal convolucional (RNC).

Como podemos apreciar en la anterior ilustración, una red neuronal convolucional 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 rectificada (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 un 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.

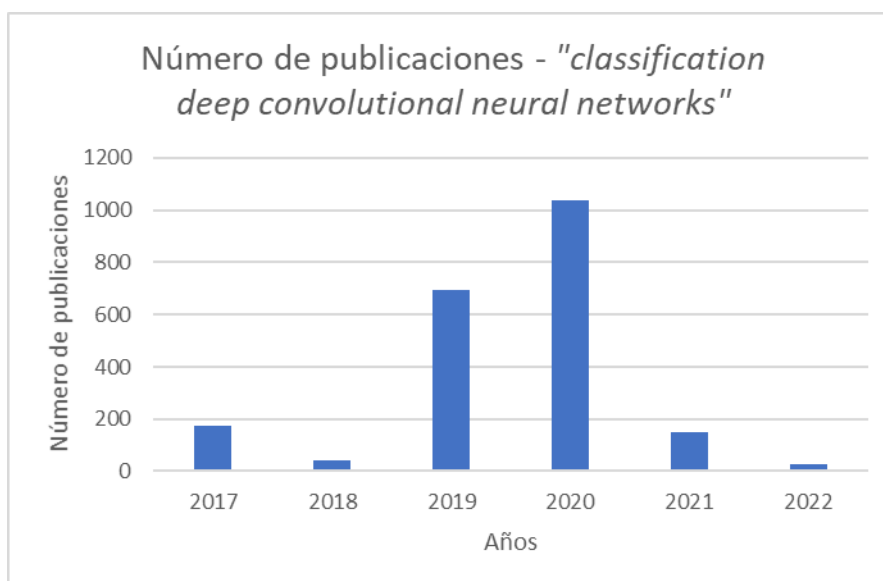


Ilustración 5 - Número de publicaciones en PubMed con la consulta "*classification deep convolutional neural networks*".

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. 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.

4. Glosario

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

5. 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 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.

6. 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.