

**DETECCIÓN DE NEUMONIA EN IMÁGENES DE RAYOS-X A PARTIR DE
PROCESAMIENTO DE IMÁGENES Y REDES NEURONALES.**

NATALIA ANDREA GÓMEZ ALBIADEZ

**UNIVERSIDAD INDUSTRIAL DE SANTANDER
FACULTAD DE INGENIERÍAS FÍSICOMECÁNICAS
ESCUELA DE INGENIERÍA DE SISTEMAS E INFORMÁTICA
BUCARAMANGA**

2021

**DETECCIÓN DE NEUMONIA EN IMÁGENES DE RAYOS-X A PARTIR DE
PROCESAMIENTO DE IMÁGENES Y REDES NEURONALES**

NATALIA ANDREA GÓMEZ ALBIADEZ

**Una tesis presentada en cumplimiento de los requisitos para el grado de
Ingeniera de Sistemas.**

Directora:

**Karen Yaneth Sánchez Quiroga
Magíster en Ingeniería Electrónica.**

Codirector:

**Henry Arguello Fuentes
Doctor en Ingeniería Eléctrica y Computación**

**UNIVERSIDAD INDUSTRIAL DE SANTANDER
FACULTAD DE INGENIERÍAS FÍSICOMECÁNICAS
ESCUELA DE INGENIERÍA DE SISTEMAS E INFORMÁTICA
BUCARAMANGA**

2021

AGRADECIMIENTOS

A mis padres, Roger Gómez y Amparo Albiadez, cada logro cumplido siempre será dedicado a ustedes, por su equilibrio entre consentimiento y disciplina, por su esfuerzo, apoyo y amor que me han dado.

A mi hermana Ginna Gómez, por siempre estar dispuesta a guiarme, quién me ha apoyado, aconsejado y actuado como una segunda madre. Siempre incondicional y orgullosa de mí, esa confianza me reconfortaba en los momentos más difíciles, siendo ella la inspiración de este y muchos logros.

A mi Abuelita Sofia quién tiene la satisfacción de verme culminar esta etapa, por brindarme su dulzura y amor en todo momento.

A Alejandro Córdoba, la persona con quién reuní experiencias personales y académicas, un amigo incondicional que sin duda hizo de este camino más llevadero, por su apoyo, compañía y aprecio.

A los ángeles que Dios me puso en el camino, y en general, mis compañeros con quién compartí a lo largo de este proceso, en tiempo de esparcimiento y trabajo. De cada uno me llevo los mejores recuerdos.

A mi directora de tesis Karen Sánchez, por su acogida, compromiso y orientación, desde el inicio de mi formación en HDSP.

Al Grupo de investigación HDSP por otorgarme los recursos necesarios para la ejecución de este proyecto.

DEDICATORIA

Este sueño recién comienza, todo es por y para ustedes.

Familia Gómez Albiadez.

CONTENIDO

	pág.
INTRODUCCIÓN	13
1. OBJETIVOS	18
1.1. OBJETIVO GENERAL	18
1.1.1. Diseñar e implementar un algoritmo de procesamiento y clasificación de imágenes de rayos-X para el apoyo al diagnóstico de neumonía mediante técnicas de transferencia de aprendizaje en redes neuronales convolucionales.	18
1.2. OBJETIVOS ESPECÍFICOS	18
1.2.1. Seleccionar una base de datos de radiografías de tórax en humanos con y sin neumonía, etiquetada por personal médico experto, disponible en internet y avalada para fines de investigación.	18
1.2.2. Diseñar una estrategia de procesamiento y clasificación de las imágenes de rayos-X con y sin neumonía a partir de redes neuronales convolucionales.	18
1.2.3. Implementar un algoritmo para ejecutar la estrategia diseñada, incorporando la técnica de transferencia de aprendizaje.	18
1.2.4. Evaluar el rendimiento del algoritmo de procesamiento y clasificación de imágenes de rayos-X para el diagnóstico de neumonía en términos de tiempo y exactitud.	18
1.2.5. Comparar los resultados obtenidos con otros métodos similares del estado del arte.	18
2. MARCO DE REFERENCIA	19

2.1. MARCO TEÓRICO	19
2.1.1. Neumonía.	19
2.1.2. Imágenes de rayos-X.	21
2.1.3. Procesamiento de imágenes.	24
2.1.4. Aprendizaje automático.	28
2.1.5. Aprendizaje profundo.	29
2.1.6. Redes neuronales convolucionales.	36
2.1.7. Transferencia de aprendizaje.	39
2.2. ESTADO DEL ARTE	41
2.3. MARCO TECNOLÓGICO	43
2.3.1. Software.	43
2.3.2. Hardware.	44
3. METODOLOGÍA	45
3.1. BASE DE DATOS	45
3.2. PROCESAMIENTO DE IMÁGENES	46
3.2.1. Aumento de datos.	46
3.2.2. Ecuilibración de histogramas.	48
3.2.3. Redimensión de imágenes.	48
3.3. MODELO DE APRENDIZAJE PROFUNDO PARA LA CLASIFICACIÓN DE NEUMONÍA	49
3.3.1. Arquitectura VGG-16.	49
3.3.2. Arquitectura ResNet-50.	50
3.3.3. Arquitectura Inception V3.	51
3.3.4. Arquitectura DenseNet-169.	52
3.4. MÉTRICAS DE VALIDACIÓN	55
3.4.1. Exactitud (<i>Accuracy</i>).	56
3.4.2. Precisión (PRE).	56

3.4.3. Exhaustividad (<i>Recall</i>).	57
3.4.4. Matriz de confusión.	57
4. RESULTADOS	59
4.1. Selección de base de datos.	59
4.2. Estrategia de Procesamiento y Clasificación	61
4.3. Pre-procesamiento.	62
4.4. Implementación del algoritmo	63
4.5. Evaluación del rendimiento del algoritmo	64
4.6. Revisión del estado del arte.	68
5. CONCLUSIONES	73
BIBLIOGRAFÍA	75
ANEXOS	86

LISTA DE FIGURAS

	pág.
Figura 1. Neumonía	20
Figura 2. Rayos-X en distintas zonas del cuerpo	22
Figura 3. Ejemplos de tecnología para diagnóstico.	25
Figura 4. Técnica de ecualización	27
Figura 5. Inteligencia Artificial	28
Figura 6. Ilustración Aprendizaje automático Vs aprendizaje profundo	29
Figura 7. Arquitectura neuronal	30
Figura 8. Proceso de aprendizaje	31
Figura 9. Clasificación multiclase-Softmax	32
Figura 10. Función de activación	34
Figura 11. Forwardpropagation y backwardpropagation	35
Figura 12. Arquitectura CNN	37
Figura 13. Subsampling	38
Figura 14. Transferencia de aprendizaje	40
Figura 15. VGG-16	49
Figura 16. DenseNet	51
Figura 17. Inception V3	52
Figura 18. DenseNet	53
Figura 19. Captura sitio web del dataset	60
Figura 20. Estrategia propuesta	61
Figura 21. Imagen ecualizada	62
Figura 22. Consolidado iteraciones CNNs	66

Figura 23. Matriz de confusión DenseNet-169 67

Figura 24. Resultados DenseNEt-169 72

LISTA DE TABLAS

	pág.
Tabla 1. Clasificación de energía en mSv	23
Tabla 2. Distribución Dataset	60
Tabla 3. Hiperparámetros CNNs	64
Tabla 4. Resultados	65
Tabla 5. Resultados cuantitativos	65
Tabla 6. Comparación con trabajos similares	68
Tabla 7. Réplica de trabajos similares	69

RESUMEN

TÍTULO: DETECCIÓN DE NEUMONIA EN IMÁGENES DE RAYOS-X A PARTIR DE PROCESAMIENTO DE IMÁGENES Y REDES NEURONALES. *

AUTOR: NATALIA ANDREA GÓMEZ ALBIADEZ **

PALABRAS CLAVE: CLASIFICACIÓN, NEUMONÍA, REDES NEURONALES CONVOLUCIONALES, PROCESAMIENTO DE IMAGEN, TRANSFERENCIA DE APRENDIZAJE.

DESCRIPCIÓN: Las infecciones respiratorias agudas son una de las principales causas de mortalidad en Bucaramanga. Además, la emergencia sanitaria producida por el virus del covid-19, ha encendido las alarmas mundiales de investigación en torno a la detección oportuna y precisa de enfermedades respiratorias agudas, lo cual es determinante para el éxito de sus tratamientos. Específicamente, una de las herramientas más usadas para realizar diagnósticos, planear tratamientos y monitorear cambios de alguna patología a través del tiempo, es la interpretación de imágenes médicas adquiridas bajo distintas modalidades. Sin embargo, además de la adquisición de las imágenes médicas se requiere de una consulta con un profesional del área de la salud para su interpretación. Este proceso en países en vía de desarrollo como Colombia suele ser lento y costoso.

Adicionalmente, se prevé que la creación de alternativas tecnológicas que faciliten el diagnóstico remoto permitirá acortar las brechas de atención médica de calidad entre zonas rurales y urbanas; y apoyará en términos de tiempo y precisión el trabajo del personal médico vinculado a los centros urbanos de salud. Por otra parte, los avances más recientes del estado de arte posicionan a la inteligencia artificial, y específicamente al aprendizaje profundo mediante redes neuronales convolucionales como una de las técnicas más prometedoras para el diagnóstico asistido por computador. Por tal motivo, se estudió una estrategia automatizada y rápida para la detección de neumonía a través de imágenes de rayos-X de tórax mediante redes neuronales convolucionales.

* Trabajo de grado

** Facultad de Ingenierías Físico-Mecánicas. Escuela de Ingeniería de Sistemas e Informática. Director: Karen Sanchez, MEng. Codirector: Henry Arguello, Ph.D.

ABSTRACT

TITLE: DETECTION OF PNEUMONIA IN X-RAY IMAGES FROM IMAGE PROCESSING AND NEURAL NETWORKS. *

AUTHOR: NATALIA ANDREA GÓMEZ ALBIADEZ. **

KEYWORDS: CLASSIFICATION, CONVOLUTIONAL NEURAL NETWORKS, IMAGE PROCESSING, PNEUMONIA, TRANSFER LEARNING.

DESCRIPTION: Acute respiratory infections are one of the leading causes of mortality in Bucaramanga, Colombia. Besides, the health emergency produced by the covid-19 virus has set off global research alarms around the timely and accurate detection of acute respiratory diseases, which is decisive for the success of their treatments. Specifically, one of the most used tools to perform diagnoses, plan treatments and monitor changes in any pathology over time is the interpretation of medical images acquired under different modalities. However, in addition to the acquisition of medical images, a consultation with a health professional is required for their interpretation. This process in developing countries like Colombia is often slow and expensive.

Additionally, It Is Expected That The Creation Of Technological Alternatives That Facilitate Remote Diagnosis Will Make It Possible To Shorten The Gaps In Quality Medical Care Between Rural And Urban Areas; and it will support in terms of time and precision the work of medical personnel linked to urban health centers. On the other hand, the most recent advances in the state of the art position artificial intelligence, and specifically deep learning through convolutional neural networks, as one of the most promising techniques for computer-assisted diagnosis. For this reason, an automated and rapid strategy was studied for the detection of pneumonia through chest X-ray images using convolutional neural networks.

* Degree work

** Faculty of Physics-Mechanics Engineering. School of Systems Engineering and Informatics. Advisor: Karen Sanchez, MEng. Co-advisor: Henry Arguello, PhD.

INTRODUCCIÓN

Las Infecciones Respiratorias Agudas (IRA) constituyen un grupo de enfermedades que se producen en el aparato respiratorio, las cuales pueden ser causadas por diferentes microorganismos como virus, bacterias u hongos.

Particularmente, la neumonía es la IRA más frecuente en el mundo ¹ y representa la principal causa de mortalidad infantil según la Organización Mundial de Salud (OMS) ². Específicamente, en el año 2015 la neumonía ocasionó el deceso de más de 900.000 niños en el mundo, es decir, uno cada 39 segundos; siendo responsable del 15 % de todas las defunciones de menores de 5 años. En el contexto local, según datos del Observatorio Digital de Datos Abiertos ³ más de 200.000 personas fueron atendidas en Bucaramanga en el año 2019 debido a IRAs, de las cuales 81 fallecieron.

En detalle, los pulmones están formados por pequeños sacos llamados alvéolos, que en las personas sanas se llenan de aire al respirar. Los alvéolos de las personas enfermas de neumonía están llenos de pus y líquido, lo que hace dolorosa la respiración y limita la absorción de oxígeno. Un diagnóstico correcto y oportuno de

¹ MINSALUD. *Infecciones Respiratorias Agudas - Ministerio de Salud y protección Social*. 2020.

² OMS. *Organización Mundial de Salud-Neumonía*. 2019.

³ DATOS.GOV.CO. *Alcaldía de Bucaramanga - Observatorio Digital Municipal de Bucaramanga-Portal de Datos Abiertos*. 2019.

la neumonía es fundamental para el éxito del tratamiento en el paciente ⁴. Generalmente, el diagnóstico se realiza a partir de la historia clínica y/o un examen físico. Aun así, expertos médicos coinciden en que el método más preciso para el diagnóstico de esta enfermedad es el análisis visual de una radiografía de tórax (RTX) adquirida con un dispositivo de rayos-X o de una tomografía computarizada (TC). Debido al costo y dosis de radiación de ambos tipos de imágenes, las RTX son usadas con mayor frecuencia para el diagnóstico de esta enfermedad ^{5,6}.

Sin embargo, completar un diagnóstico a través del análisis visual de estas imágenes especializadas implica un lapso entre radiólogo y médico que impide al paciente el acceso temprano a su diagnóstico. Adicionalmente, en países en vía de desarrollo como Colombia, el acceso oportuno a este tipo de exámenes especializados es escaso, especialmente en zonas rurales, y generalmente su interpretación toma un tiempo adicional considerable. Como si fuera poco, múltiples estudios afirman que el diagnóstico de enfermedades a partir de interpretación visual, aún cuando es realizado por personal médico especializado, está propenso a variaciones entre uno y otro profesional, en términos de dictamen de la presencia o ausencia de enfermedad, y nivel de gravedad de la misma ⁷.

⁴ TAKEMBO, Clovis Ntahkie, *et al.* "Effect of electromagnetic radiation on the dynamics of spatio-temporal patterns in memristor-based neuronal network". En: *Nonlinear Dynamics* 95.2 (2019), págs. 1067-1078.

⁵ HRESCAK, María Cinthya Ortega y SOCOLSKY, Gustavo A. "Godfrey Newbold Hounsfield: historia e impacto de la tomografía computada". En: *Revista argentina de radiología* 76.4 (2012), págs. 331-341.

⁶ ESTEVAN, Miguel. "Examen radiográfico del tórax en las neumonías de probable causa bacteriana". En: *Archivos de Pediatría del Uruguay* 73.1 (2002), págs. 15-21.

⁷ LITJENS, Geert, *et al.* "A survey on deep learning in medical image analysis". En: *Medical image analysis* 42 (2017), págs. 60-88.

Por tal motivo, el diagnóstico asistido por computador (en inglés *computer-aided diagnosis*) (CAD) se ha convertido en un área de investigación activa y desafiante^{8,9,10}. El propósito del CAD es otorgar una herramienta de soporte al radiólogo mediante la interpretación de imágenes médicas a partir de inteligencia artificial.

Recientemente, se han propuesto diversas soluciones basadas en aprendizaje profundo (en inglés *Deep Learning*) para diferentes problemas médicos, como la detección del cáncer de mama, predicción de enfermedades como, Alzheimer, clasificación de enfermedades en imágenes de rayos-X¹¹, entre otros. Además los avances más recientes del estado de arte han demostrado que corregir y destacar características de las imágenes a través de operaciones de procesamiento, previo a alimentar la red neuronal convolucional, permite incrementar la precisión y rendimiento final de

-
- ⁸ KERMANY, Daniel S, *et al.* "Identifying medical diagnoses and treatable diseases by image-based deep learning". En: *Cell* 172.5 (2018), págs. 1122-1131.
- ⁹ GAMAGE, C., *et al.* "GI-Net: Anomalies Classification in Gastrointestinal Tract through Endoscopic Imagery with Deep Learning". En: *2019 Moratuwa Engineering Research Conference (MER-Con)*. 2019, págs. 66-71.
- ¹⁰ RAJPURKAR, Pranav, *et al.* "Chexnet: Radiologist-level pneumonia detection on chest x-rays with deep learning". En: *arXiv preprint arXiv:1711.05225* (2017).
- ¹¹ KHAN, SanaUllah, *et al.* "A novel deep learning based framework for the detection and classification of breast cancer using transfer learning". En: *Pattern Recognition Letters* 125 (2019), págs. 1-6.

la clasificación^{12,13,14}. Así mismo, han demostrado resultados significativos utilizando filtros convolucionales aplicados en redes neuronales de aprendizaje profundo para la detección de estructuras y características sospechosas, que generalmente podrían ser omitidas por los profesionales de la salud mediante inspección visual¹⁵.

En particular, existe la urgente necesidad de desarrollar y/o mejorar herramientas de detección automatizada de neumonía en imágenes de rayos-X de tórax, inmediatamente después de su adquisición. Una petición que respaldan los profesionales de la salud que se han visto enfrentados a la interpretación de gran cantidad de radiografías de tórax para el diagnóstico de Covid-19, la más reciente emergencia sanitaria mundial cuyo principal síntoma es la afectación aguda de las vías respiratorias^{16,17}.

Las redes neuronales convolucionales de aprendizaje profundo se caracterizan por requerir una gran cantidad de datos etiquetados iniciales para entrenar el modelo de clasificación, lo cual genera problemas que van desde los relacionados con la

¹² AGUIRRE DOBERNACK, Nicolás. "Implementación De Un Sistema De Detección De Señales De Tráfico Mediante Visión Artificial Basado En FPGA". En: *Universidad de Sevilla* (2013).

¹³ PÉREZ, P. y VALENTE, M. *Fundamentos básicos del procesamiento de imágenes*. 2018.

¹⁴ FERNÁNDEZ, Luis A; DIAZ, Daniel y DEPAOLI, Roberto. "Optimización de la ecualización del histograma en el procesamiento de imágenes digitales". En: *VII Workshop de Investigadores en Ciencias de la Computación*. 2005.

¹⁵ ASNAOUI, Khalid El; CHAWKI, Youness e IDRI, Ali. "Automated methods for detection and classification pneumonia based on x-ray images using deep learning". En: *arXiv preprint arXiv:2003.14363* (2020).

¹⁶ SONG, Fengxiang, *et al.* "Emerging 2019 novel coronavirus (2019-nCoV) pneumonia". En: *Radiology* 295.1 (2020), págs. 210-217.

¹⁷ GATTINONI, Luciano, *et al.* *COVID-19 pneumonia: different respiratory treatments for different phenotypes?* 2020.

privacidad del paciente hasta la gestión y almacenamiento de los datos, resultando costoso y complejo¹⁸. Por tal motivo, la transferencia de aprendizaje se ha convertido en una rama de alto interés para el diagnóstico médico asistido por computador, además se considera como una de las técnicas más importantes del aprendizaje profundo en inteligencia artificial, ya que permite el desarrollo de tareas de clasificación, a partir de redes neuronales previamente entrenadas con otra base de datos, lo que implica un recurso economizador de tiempo, costo computacional y cantidad de imágenes médicas^{19,20}.

En este trabajo se utiliza un enfoque de transferencia de aprendizaje basado en Redes Neuronales Convolucionales (CNN), se implementó un algoritmo de procesamiento y clasificación de imágenes de rayos-X de tórax de vista frontal, para el apoyo al diagnóstico de neumonía utilizando una arquitectura previamente entrenada en ImageNet, usamos tres modelos diferentes previamente entrenados y analizamos su desempeño.

¹⁸ PAN, Sinno Jialin y YANG, Qiang. "A survey on transfer learning". En: *IEEE Transactions on knowledge and data engineering* 22.10 (2009), págs. 1345-1359.

¹⁹ RAHMAN, Tawsifur, *et al.* "Transfer Learning with Deep Convolutional Neural Network (CNN) for Pneumonia Detection using Chest X-ray". En: *Applied Sciences* 10.9 (2020), pág. 3233.

²⁰ CHOUHAN, Vikash, *et al.* "A novel transfer learning based approach for pneumonia detection in chest X-ray images". En: *Applied Sciences* 10.2 (2020), pág. 559.

1. OBJETIVOS

1.1. OBJETIVO GENERAL

1.1.1. Diseñar e implementar un algoritmo de procesamiento y clasificación de imágenes de rayos-X para el apoyo al diagnóstico de neumonía mediante técnicas de transferencia de aprendizaje en redes neuronales convolucionales.

1.2. OBJETIVOS ESPECÍFICOS

1.2.1. Seleccionar una base de datos de radiografías de tórax en humanos con y sin neumonía, etiquetada por personal médico experto, disponible en internet y avalada para fines de investigación.

1.2.2. Diseñar una estrategia de procesamiento y clasificación de las imágenes de rayos-X con y sin neumonía a partir de redes neuronales convolucionales.

1.2.3. Implementar un algoritmo para ejecutar la estrategia diseñada, incorporando la técnica de transferencia de aprendizaje.

1.2.4. Evaluar el rendimiento del algoritmo de procesamiento y clasificación de imágenes de rayos-X para el diagnóstico de neumonía en términos de tiempo y exactitud.

1.2.5. Comparar los resultados obtenidos con otros métodos similares del estado del arte.

2. MARCO DE REFERENCIA

2.1. MARCO TEÓRICO

2.1.1. Neumonía. La neumonía se identifica por ser una infección que se presenta en la zona de los pulmones y es la enfermedad más frecuente entre los pacientes con insuficiencia respiratoria aguda. Su origen puede estar asociado a virus o a bacterias, y debido a las graves complicaciones que puede generar en los pacientes y a la necesidad de ventilación mecánica como medida de asistencia a la respiración, su tratamiento clínico representa un alto costo para las empresas prestadoras de servicios médicos. La principal vía de entrada de los patógenos al organismo humano es por vía aérea, a través de la aspiración de secreciones orofaríngeas. De esta manera, microbacterias, hongos, microorganismos o virus ingresan a los pulmones, se alojan y desencadenan infecciones que provocan que en los sacos de aire se llene de material purulento²¹. Esta infección puede ser mortal para cualquier persona, pero, especialmente para bebés, niños y adultos mayores de 65 años²². En el 2015 la neumonía fue la causa de muerte de 920.000 niños en todo el mundo, y actualmente, representa la tercera causa de fallecimiento en Bucaramanga, Colombia. Por otra parte, las técnicas para el diagnóstico de la enfermedad de neumonía suelen ser:

- Análisis de sangre para confirmar una infección e identificar el tipo de organismo que la genera.

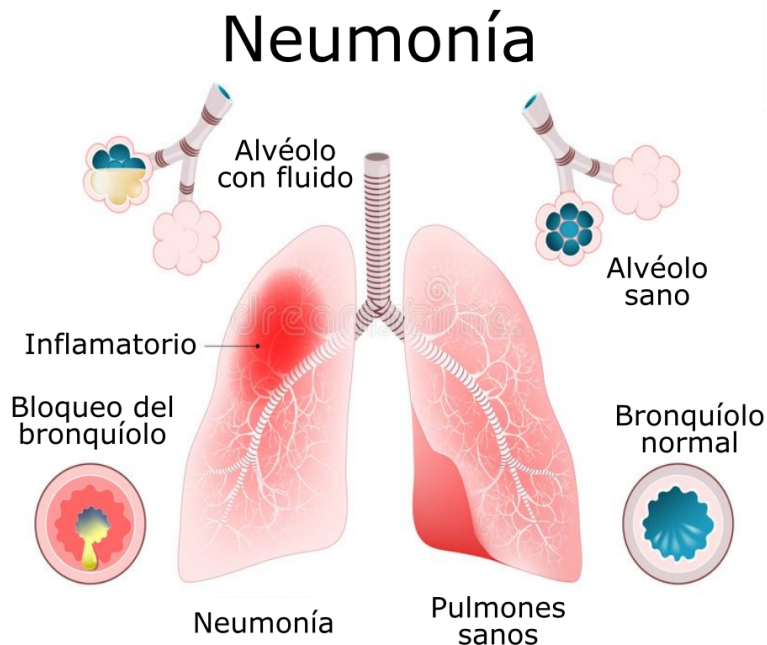
²¹ DÍAZ, Emili; MARTÍN-LOECHES, Ignacio y VALLÉS, Jordi. "Neumonía nosocomial". En: *Enfermedades Infecciosas y Microbiología Clínica* 31.10 (2013), págs. 692-698.

²² LOVERA, Dolores, *et al.* "Neumonía en lactantes de 6 meses de edad y menores". En: *Pediatría (Asunción)* 47.2 (2020), págs. 74-80.

- Radiografía torácica.
- Pulsioximetría.
- Prueba de esputo.

En particular, las radiografías son una de las técnicas más utilizadas en el mundo para el diagnóstico de esta enfermedad.

Figura 1. El pulmón izquierdo hace referencia clínica a la neumonía, en donde se puede evidenciar los alvéolos llenándose de fluidos y pus (líquido amarillo), a diferencia del pulmón derecho que se encuentra sano. Tomado de²³.



De manera que, si una persona presenta los síntomas característicos, que incluyen tos con flema o pus, fiebre, escalofríos y dificultad para respirar, es remitida a un radiólogo que captura una imagen de la zona torácica del paciente con un dispositivo de rayos-X. Posteriormente, un médico experto interpreta la imagen, y determina si corresponde o no a un caso de neumonía. Para el diagnóstico, el experto analiza en la imagen si existe presencia de consolidación de espacio aéreo periférico,

que traduce la sustitución de aire por exudado, dando la llamada imagen acinar, representada en color blanco en la radiografía. Sin embargo, la interpretación precisa de estas imágenes requiere de años de formación, entrenamiento y experiencia, y puede variar entre uno y otro profesional.

2.1.2. Imágenes de rayos-X. Los rayos-X son un tipo de radiación electromagnética, invisible para el ojo humano, capaz de atravesar cuerpos opacos y generar una imagen en el computador o en una película fotográfica²⁴. Específicamente, los rayos-X se utilizan en el área de la salud para generar imágenes de los huesos y tejidos del interior del cuerpo humano, de la siguiente manera:

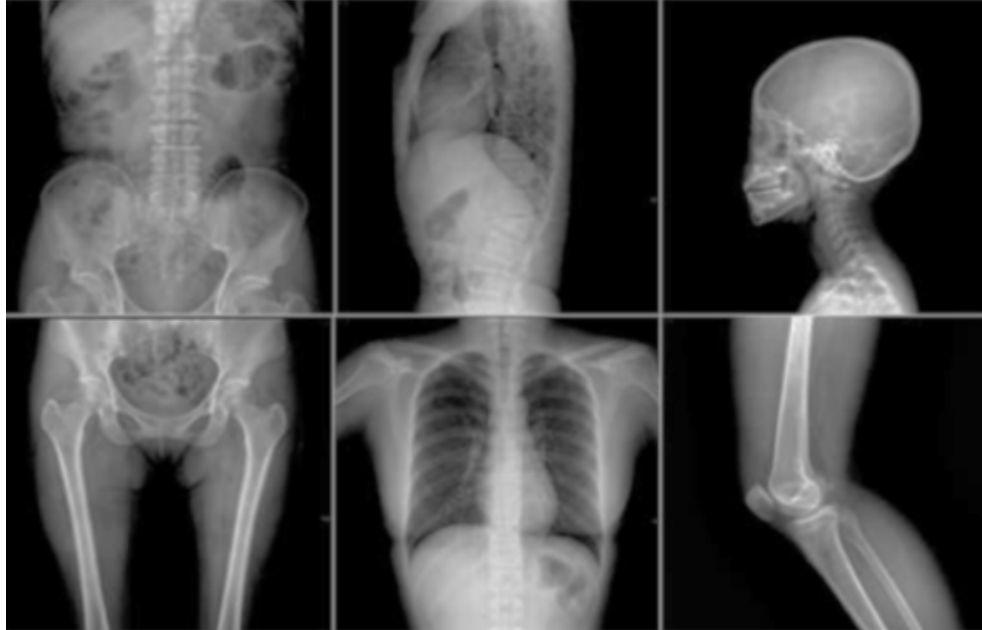
- Las estructuras óseas, el metal y los medios de contraste, debido a su alta densidad bloquearán la mayoría de las partículas de rayos-X, por lo que estas zonas serán de color blanco en la imagen.
- Tejidos como los músculos, la grasa y los líquidos aparecerán en tonos de color gris.
- El aire contenido, debido a su baja densidad que permite el paso de la radiación, aparecerá en color negro.

Una imagen de rayos-X es interpretada computacionalmente como una matriz de valores numéricos entre 0 y 255, con un número por pixel, donde el valor 0 representa el color negro y 255 el blanco²⁵. En la figura-2 se observan imágenes típicas de rayos-X de diferentes zonas del cuerpo humano.

²⁴ CHUNG, Albert Chi Shing, *et al.* "Information processing in medical imaging". En: *Information Processing in Medical Imaging*. 2019.

²⁵ SAHINER, Berkman, *et al.* "Deep learning in medical imaging and radiation therapy". En: *Medical physics* 46.1 (2019), e1-e36.

Figura 2. Ilustración de radiografías en distintas zonas del cuerpo como, abdomen frontal, abdomen lateral, cráneo lateral, pelvis, tórax y rodilla. Tomado de²⁶.



Por otra parte, cada vez que se adquiere una imagen de rayos-X el área de exposición está directamente relacionada con la dosis de radiación recibida por el paciente. Esta cantidad de energía, denominada dosis efectiva, es cuantificada en milisievert (mSv)²⁷ tal como lo expone la tabla-1. Dicha radiación a alta velocidad puede inferir en el comportamiento natural de la energía ionizante al interior del cuerpo, razón por la cual, una alteración significativa del estado natural podría llegar a incrementar el riesgo de desencadenar cáncer en el futuro; sin embargo, la probabilidad de aumento de riesgo por una sola imagen convencional es muy baja ²⁸.

²⁷ NORTH AMERICA (RSNA), Radiological Society of. *Seguridad del paciente: Dosis de radiación en exámenes de rayos X y TAC Qué son los rayos X y qué hacen.* 2013.

²⁸ WILLEMINK, Martin J, *et al.* "Preparing medical imaging data for machine learning". En: *Radio-logy* 295.1 (2020), págs. 4-15.

Por otra parte, la energía electromagnética transmitida por distintas antenas, estaciones y dispositivos electrónicos que hacen parte del entorno de hoy, han desarrollado una radiación permanente que se conoce como radiación natural, la cual también es comparada en la tabla con la energía proporcionada por la adquisición de imágenes de rayos-X ^{5,29}.

Tabla 1. Dosis recibida en diferentes procedimientos diagnósticos a un número similar de radiografías de tórax y radiación por fuentes naturales. Adaptada de ³⁰.

Procedimiento Diagnóstico	Dosis efectiva (mSv)	Incremento riesgo de cáncer	Radiación natural
Radiología			
Tórax	0.02	1 en 1.000.000	3 días
Extremidades	0.01	1 en 1.000.000	1.5 días
Cráneo	0.07	1 en 10.000	11 días
Columna Dorsal	0.70	1 en 10.000	4 meses
Columna Lumbar	1.30	1 en 10.000	7 meses
Abdomen	1.00	1 en 10.000	6 meses
Pelvis	0.70	1 en 10.000	4 meses
Urografía	2.50	1 en 10.000	14 meses
Rx EED	3.00	1 en 10.000	16 meses
Enema opaco	7.00	1 en 10.000	3.2 años
Tc de Cráneo	2.30	1 en 2.000	1 años
Tc de Tórax	8.00	1 en 2.000	3.6 años
Tc de abdomen	10.0	1 en 2.000	4.5 años

Las primeras imágenes de rayos-X fueron obtenidas por Godfrey Newbold Hounsfield mediante un escáner basado en técnicas algebraicas, fundadas en reconocimientos de patrones determinados por las variables: ancho del haz de irradiación,

²⁹ URIBARRI, Salvador Galindo. "Principios matemáticos de la reconstrucción de imágenes tomográficas". En: *CIENCIA ergo-sum* 10.3 (2003), págs. 271-281.

ángulo de irradiación y tamaño de sección del objeto a reconstruir^{31,32}. Actualmente, se mantienen los principios algebraicos de estos sistemas de adquisición; sin embargo, algunos dispositivos han pasado de usar películas fotográficas a adquirir las imágenes digitalmente para ser visualizadas en una pantalla de computador³³. Las imágenes que se planea utilizar en el presente proyecto provendrán de una base de datos de radiografías disponibles digitalmente en internet.

2.1.3. Procesamiento de imágenes. Una imagen puede ser definida como una matriz de tamaño m filas por n columnas, donde $f(x, y)$ es el valor de intensidad de cada pixel en la coordenada (x, y) . Una imagen se configura como una sucesión de píxeles, ya que el grado de detalle o calidad de una imagen está determinado por el tamaño del píxel, a un tamaño más pequeño y a un mayor número de píxeles se adquiere una imagen con alta resolución espacial¹². Para una imagen a escala de grises el valor de cada píxel estará en el rango $[0, 255]$ siendo 0 la representación del color negro y 255 blanco. Además, cada imagen contiene información espectral a una longitud de onda dada que consiste en el número de canales espectrales conocidos como la banda espectral. Según la resolución espacio-espectral, las imágenes se clasifican en pancromáticas (escalas de grises) las cuales contienen una banda espectral, las RGB con tres bandas espectrales (*Red, Green, Blue*), multi-espectrales e hiperespectrales con cientos y hasta miles de bandas espectrales,

³¹ WORLDHEALTHORGANIZATION, *et al.* *Standardization of interpretation of chest radiographs for the diagnosis of pneumonia in children*. Inf. téc. World Health Organization, 2001.

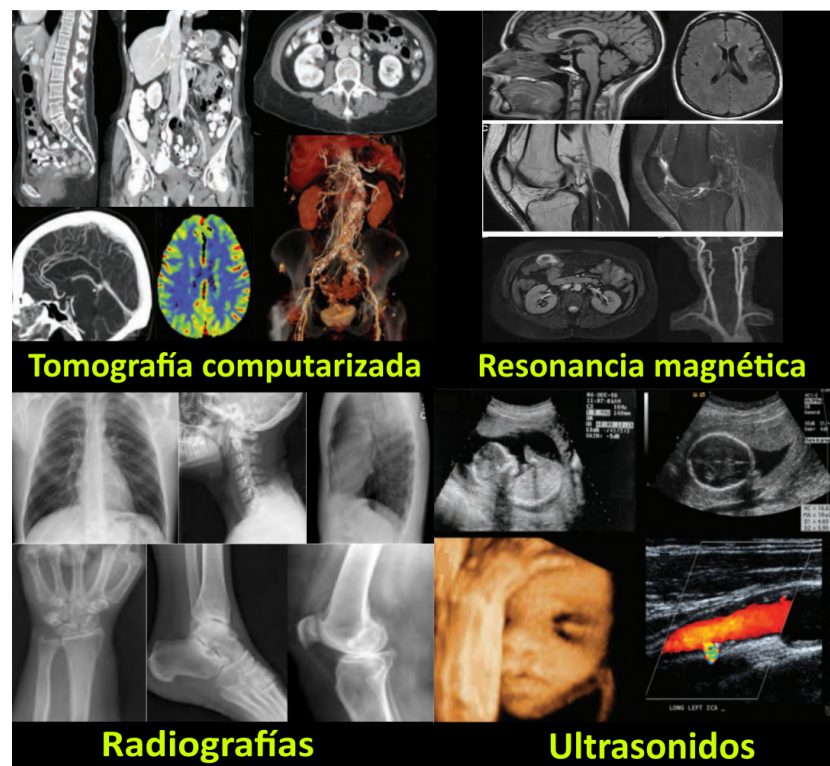
³² KETCHAM, Richard A y CARLSON, William D. "Acquisition, optimization and interpretation of X-ray computed tomographic imagery: applications to the geosciences". En: *Computers & Geosciences* 27.4 (2001), págs. 381-400.

³³ KADOYA, Noriyuki, *et al.* "Evaluation of four-dimensional cone beam computed tomography ventilation images acquired with two different linear accelerators at various gantry speeds using a deformable lung phantom". En: *Physica Medica* 77 (2020), págs. 75-83.

respectivamente.

El tratamiento de imágenes es el conjunto de técnicas usadas para resaltar, mejorar, ignorar o descubrir las características de una imagen de interés mediante la mejor resolución propia de cada una de ellas³⁴.

Figura 3. Distintas tecnologías para diagnosticar condiciones médicas mediante imágenes. Adaptada de ³⁵.



El procesamiento se puede llevar a cabo mediante métodos ópticos o digitales. Específicamente, el procesamiento digital de imágenes aborda diversas técnicas que se aplican a las imágenes con el fin de mejorar su calidad resaltando detalles que

³⁴ GELVES MENDOZA, Sergio Alexander y RINCON CADENA, Wanda Catalina. "Metodos De Machine Learning Para El Apoyo Al Diagnostico De Supervivencia De Pacientes Con Sepsis Basadas En Datos Clinicos Y Zimografias." Tesis doct. Universidad Industrial de Santander, Escuela De Ing. De Sistemas, 2015.

sean convenientes, alterando los valores de la luminosidad de los píxeles mediante las operaciones o transformaciones matemáticas necesarias ³⁶. Por ejemplo, sobre la vecindad de cada pixel se le aplica un kernel con el fin de destacar ciertas característica, dicha técnica se le conoce como filtrado de imágenes. De manera específica estas técnicas de procesamiento favorecen la precisión en las aplicaciones de detección, en áreas como la medicina. La figura-3 presenta algunas de las modalidades de imágenes médicas más conocidas en la actualidad. Para la representación digital de las imágenes se cuenta con distintos métodos como mapas de Bits (Bitmaps), imágenes vectoriales, histograma de una imagen, entre otros ¹³.

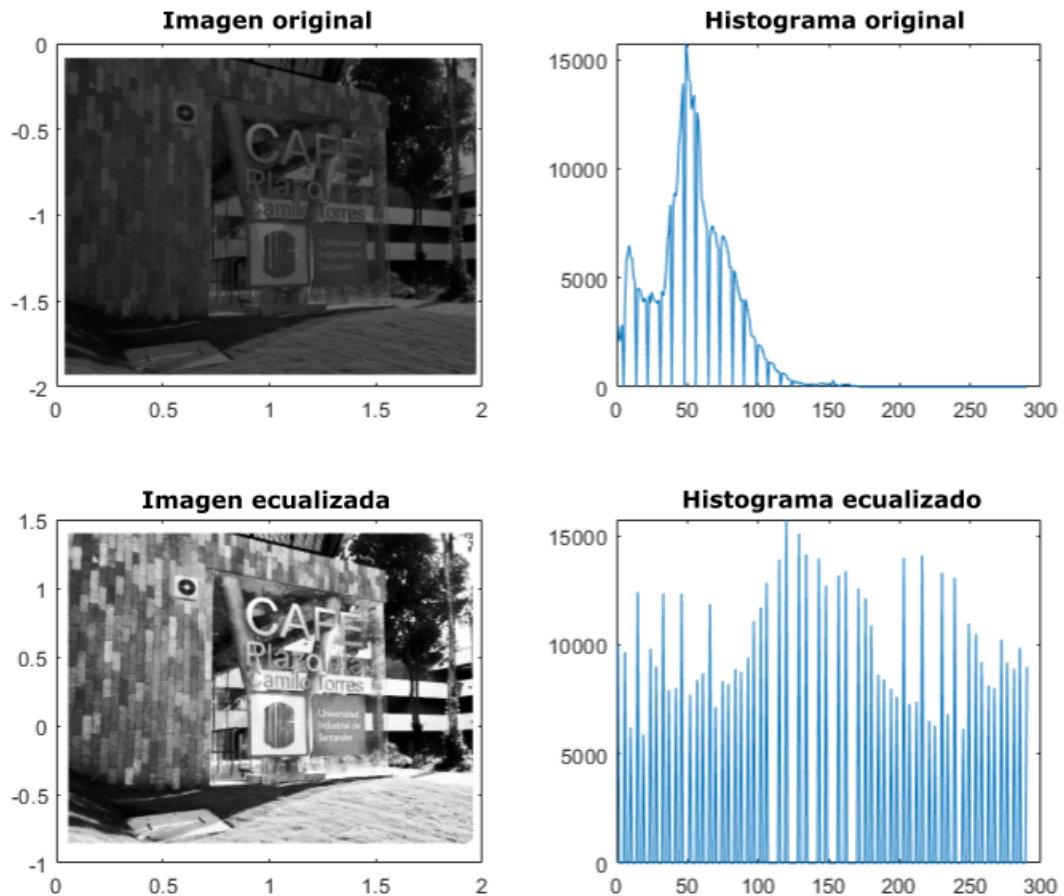
En el procesamiento de imágenes un histograma puede ser utilizado para visualizar el rango de niveles que toma la escena, en específico un histograma representa los valores de los píxeles de una imagen. El análisis estadístico derivado del histograma puede servir para comparar contrastes e intensidades dentro de los niveles de grises, su principal función es aumentar el rango dinámico de una imagen digital o realizar una ecualización, cuando sus niveles se encuentran concentrados en una zona del intervalo ³⁷.

Una de las técnicas de procesamiento de imágenes tradicionales es la Ecualización del histograma. Los histogramas son interpretaciones gráficas de datos agrupados con el fin de obtener su distribución representado por un diagrama de barras, representado por $p_k = \frac{n_k}{n}$, donde n_k es igual a la cantidad de pixeles de nivel gris k , y n corresponde a la cantidad total de pixeles ¹⁴.

³⁶ MALACARA, Daniel. *Óptica tradicional y moderna*. Sirsi) i9789681666187. 2002.

³⁷ INTERNET. *Introducción a las imágenes digitales*. 2018.

Figura 4. Técnica de procesamiento de imágenes (ecualización) aplicada a una imagen que se tomó dentro del campus de la Universidad Industrial de Santander. En la parte superior imagen original, en la parte inferior imagen ecualizada, donde se evidencia mejora en el contraste y distribución de los bits en el histograma.



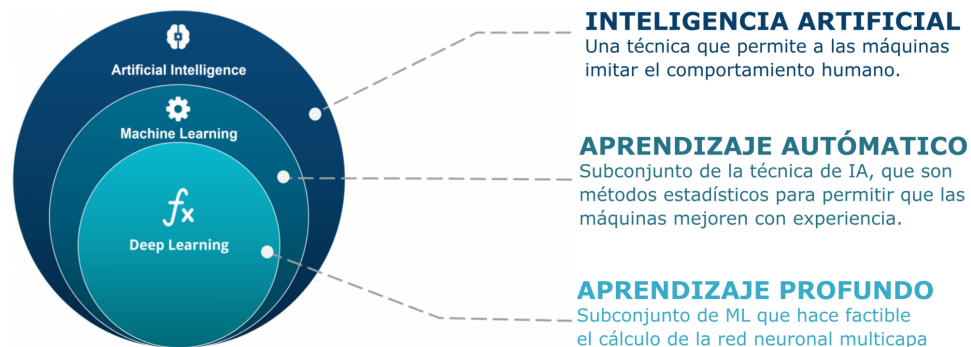
La ecualización de histogramas consiste en realizar transformaciones sobre los niveles de la imagen inspirada en una operación sobre funciones de dominio continuo, logrando una distribución de probabilidad uniforme.

Esta técnica puede ser construida mediante la normalización del histograma acumulado al rango de niveles de intensidad de la imagen, se puede aplicar directamente a imágenes en escala de grises con el fin de intensificar el contraste visual de la imagen como se muestra en la figura-4, donde se muestra una imagen sin procesar

(arriba) y una imagen cuyo histograma ha sido ecualizado (abajo). En particular, este proceso es ampliamente utilizado en radiología, debido a que maximiza el contraste de una imagen sin perder información de tipo estructural.

2.1.4. Aprendizaje automático. Aprendizaje automático (en inglés, *Machine Learning*)^{38, 39}, es un subcampo de las ciencias de la computación y una disciplina dentro del ámbito de la Inteligencia artificial, como se muestra en la figura-5.

Figura 5. La imagen ilustra los elementos que contiene el conjunto de inteligencia artificial, en general, dentro del campo del aprendizaje automático específicamente, dentro del aprendizaje supervisado, encontramos el área de aprendizaje profundo o deep learning. Adaptada de⁴⁰.



El aprendizaje automático está basado en enfoques algorítmicos programados para detectar patrones que sean utilizados como fundamentos de predicción para futuros comportamientos. Algunos modelos de aprendizaje automático son: redes neuronales artificiales, árboles de decisión, máquinas de soporte vectorial, redes Bayesianas y algoritmos genéticos. Estos modelos han sido usados durante las últimas décadas en una gran cantidad de aplicaciones, incluidas diagnóstico asistido por computador,

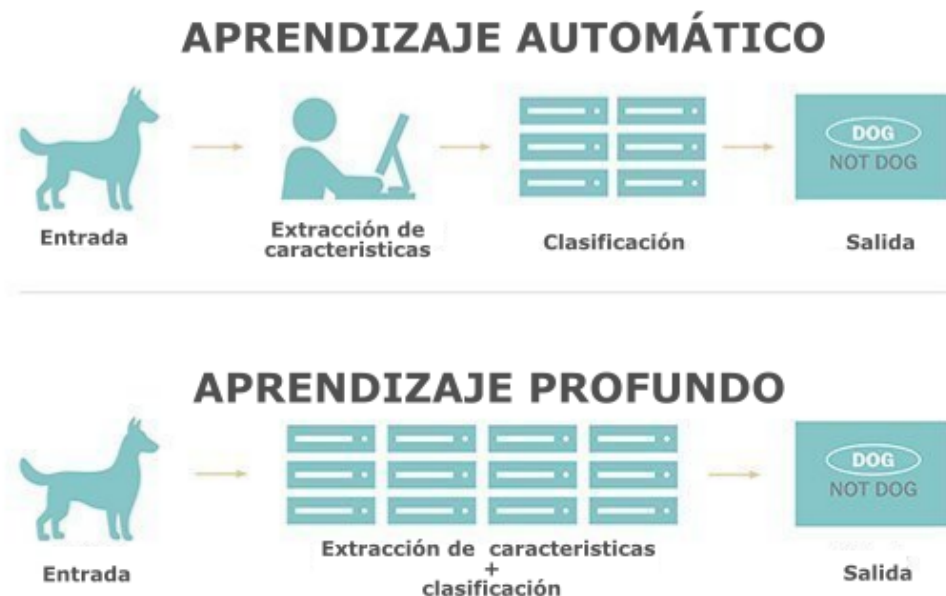
³⁸ MITCHELL, Tom M, et al. *Machine learning*. 1997.

³⁹ ALPAYDIN, Ethem. *Introduction to machine learning*. MIT press, 2020.

reconocimiento de rostros, control de plagas en cultivos, vigilancia, diseño de dispositivos ópticos, entre otros ⁴¹. Actualmente, se utilizan dos tipos principales de algoritmos de aprendizaje autónomo: aprendizaje supervisado y aprendizaje no supervisado.

2.1.5. Aprendizaje profundo. A diferencia del aprendizaje de máquina, el aprendizaje profundo no requiere de intervención humana para extraer conocimientos y suministrar cientos o miles de ejemplos a la máquina. Tal como se muestra en la figura-6, el programa construye el conjunto de características por sí mismo sin supervisión.

Figura 6. Diferencia visual entre aprendizaje automático y aprendizaje profundo. Adaptada de⁴²



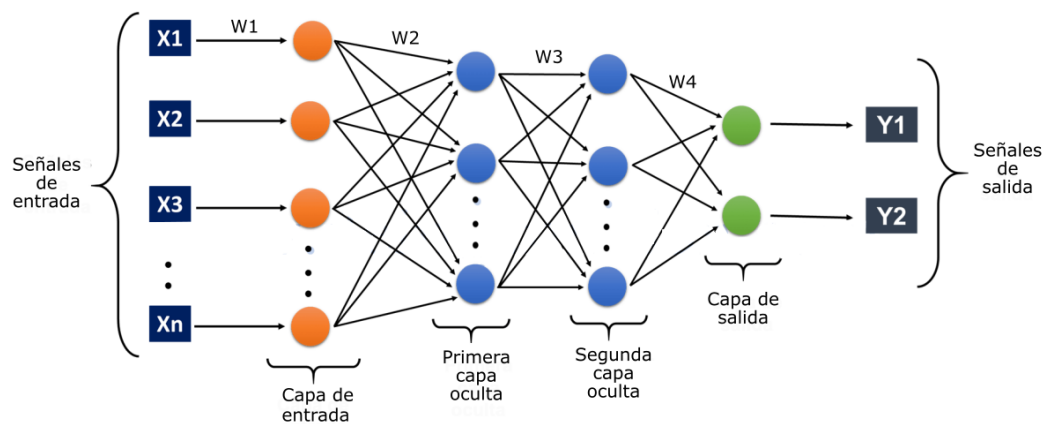
El aprendizaje profundo es un modelo computacional inspirado en el comportamiento neuronal humano, debido a que está constituido por capas anidadas en las que

⁴¹ AKHTAR, Naveed y MIAN, Ajmal. "Threat of adversarial attacks on deep learning in computer vision: A survey". En: *IEEE Access* 6 (2018), págs. 14410-14430.

se pueden encontrar con nodos interconectados y a su vez cada nodo, suministra un valor a su salida, sin embargo, las particularidades de su funcionamiento no tienen nada que ver con el funcionamiento del cerebro biológico⁴³.

La figura-7 describe una red neuronal artificial tradicional, representada por tres tipos de capas de neuronas fundamentales: capa de entrada, capa oculta y capa de salida. El tamaño de las redes depende del número de capas y del número de neuronas ocultas por capa, las redes neuronales profundas se identifican por contener numerosas capas ocultas, asimila un alto nivel de datos y se especializa en el reconocimiento de imágenes. Cuando la salida presenta dos o más neuronas, representa un problema de clasificación multiclase⁴⁴.

Figura 7. Representación visual de una arquitectura neuronal con dos capas ocultas. Adaptada de ⁴⁵.



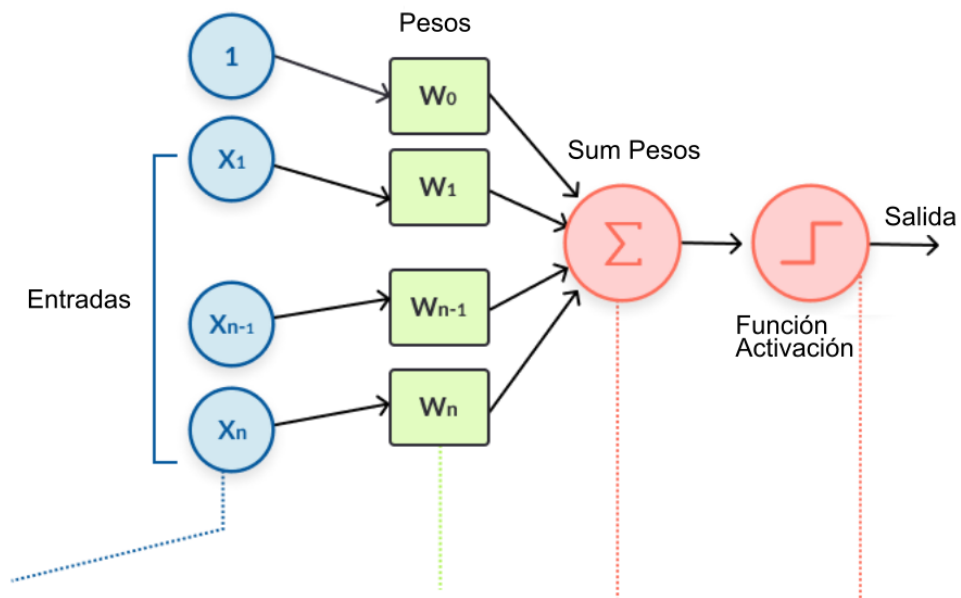
Como se mencionó anteriormente, las redes se caracterizan por contener una capa de entrada, seguido de una o más capas ocultas y una capa de salida. La capa de

⁴³ LECUN, Yann; BENGIO, Yoshua e HINTON, Geoffrey. "Deep learning". En: *nature* 521.7553 (2015), págs. 436-444.

⁴⁴ LIANG, Xuezhi, *et al.* "Soft-margin softmax for deep classification". En: *International Conference on Neural Information Processing*. Springer. 2017, págs. 413-421.

entrada está compuesta por neuronas que reciben datos o información nueva, estos datos se propagan a través de la red mediante conexiones unidireccionales hacia otros nodos de la red. A cada conexión hay asociada un peso específico-aleatorio denominado w_i , este peso define la fuerza de conexión que determina la importancia que tendrá esa relación en la neurona. Dichos pesos pueden tomar valores positivos, negativos o cero. En caso de una entrada positiva, un peso positivo actúa como excitador, mientras que un peso negativo actúa como inhibidor; si el valor es cero significa que no existe comunicación entre el par de neuronas⁴⁶.

Figura 8. La imagen describe visualmente el proceso de aprendizaje de una red, llamado propagación hacia adelante, que consiste en asignar un valor de peso a la conexión de entrada, luego realizar la suma ponderada de las entradas y pasar el resultado a una función de activación para generar una salida. Adaptada de ⁴⁷

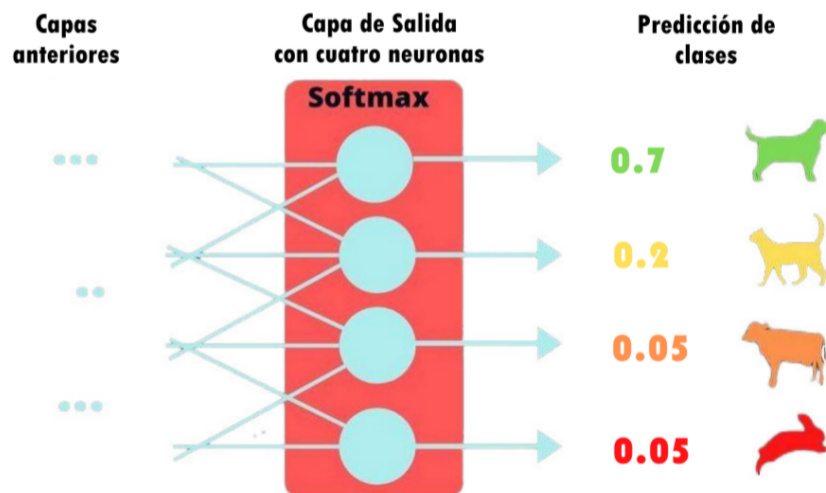


Esta información se transmite mediante una función base, transformada mediante

⁴⁶ ARELI, Jamie y BARRERA, Toral. "Redes Neuronales". [Internet; descargado 24-noviembre-2020].

una función de activación no lineal y a su vez representa la salida de la neurona como se muestra en la figura-8. Una capa de salida se compone de neuronas que proporcionan la respuesta de la red neuronal^{48,49}. Algunos ejemplos de funciones de activación son la función sigmoideal ó sigmoid, el valor dado por la función sigmoideal, es cercano a uno de los valores asintóticos. Esto hace que en la mayoría de los casos el valor de salida este comprendido en la zona alta o baja del sigmoide; la importancia de la función sigmoideal es que su derivada es siempre positiva y cercana a cero para los valores grandes positivos o negativos. Otra de las funciones frecuentemente empleadas en la capa de salida para problemas de clasificación es la función de activación softmax.

Figura 9. Distribución de pesos en una red neuronal para una clasificación no binaria. Tomada de ⁵⁰



La función de activación softmax toma N valores reales de salida y los normaliza en

⁴⁸ LÓPEZ, Raquel Florez; FERNANDEZ, Jose Miguel y FERNÁNDEZ, José Miguel Fernández. *Las redes neuronales artificiales*. Netbiblo, 2008.

⁴⁹ IZAURIETA, Fernando y SAAVEDRA, Carlos. "Redes neuronales artificiales". En: *Departamento de Física, Universidad de Concepción Chile* (2000).

una distribución de probabilidad sobre las N clases, lo cual significa que cada valor de salida estará en el intervalo de $[0, 1]$ y la suma de los N será 1 ⁵¹, tal como se muestra en la figura-9.

Por otra parte, en la figura-10 se muestran algunas de las funciones de activación más utilizadas en los distintos modelos de redes neuronales.

1. **Función sigmoide:** Tiene un gradiente suave y genera valores entre cero y uno. Para valores muy altos o bajos de los parámetros de entrada, la red puede ser muy lenta para alcanzar una predicción, llamada problema de gradiente de fuga.
2. **Función TanH:** Está centrada en cero, lo que facilita el modelado de entradas que son fuertemente negativas, fuertemente positivas o neutrales.
3. **Función ReLu:** Es altamente eficiente desde el punto de vista computacional, pero no puede procesar entradas que se acerquen a cero o negativas.
4. **Función Leaky ReLu:** Tiene una pequeña pendiente positiva en su área negativa, lo que le permite procesar valores cero o negativos.

Una función base puede ser de la forma $y = \sum_i (w_i x_i + b)$. Donde x representa el valor de entrada y b un sesgo. Este valor es determinado según el tipo de tarea a realizar. A su vez, el proceso de entrenamiento es necesario para la actualización de los pesos, esta actualización es conocida como, propagación hacia atrás (en inglés *backpropagation*), el algoritmo inicia con un valor de error aleatorio y se espera que converja a un mínimo de forma óptima mediante la actualización de sus pesos. La

⁵¹ FUENTE CASTILLO, Víctor de la y LOPE ASIAÍN, Javier de. "Diseño evolutivo de arquitecturas de Deep Learning para detección de vías de transporte." En: ().

Figura 10. En la imagen se muestran algunas de las funciones de activación más utilizadas. Adaptada de ⁵²

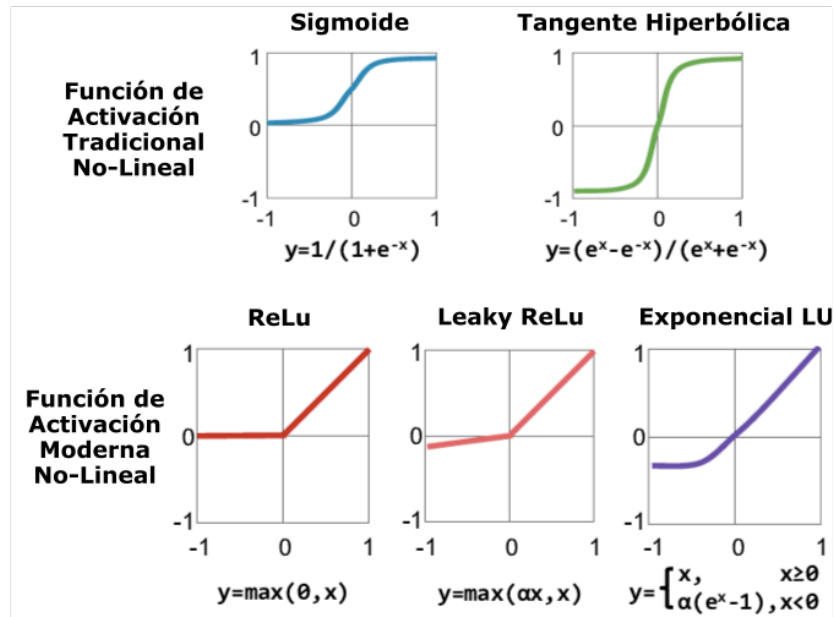


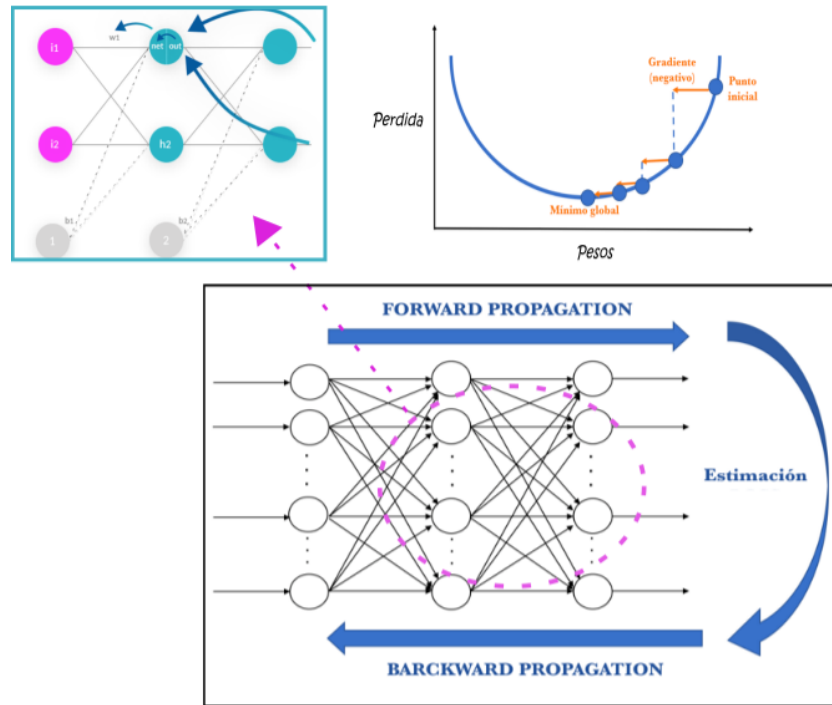
figura-11 muestra visualmente dicho proceso. Además, se requiere que las funciones de transferencia de las neuronas sean derivables ya que solamente necesita de dos ecuaciones para propagar las señales de error hacia atrás, para obtener esas ecuaciones se utilizarán tres técnicas importantes en matemáticas que es la función de error, gradiente descendiente y la regla de la cadena.

De esta manera se genera una señal de error para cada una de las salidas. Este proceso se repite capa por capa hasta que todas las neuronas de la red reciban la actualización del error relativo en contribución al error total^{54, 55}. Específicamente, dicho error se puede expresar mediante el gradiente descendiente como se mues-

⁵⁴ LARRANAGA, Pedro; INZA, Inaki y MOUJAHID, Abdelmalik. "Tema 8. redes neuronales". En: *Redes Neuronales, U. del P. Vasco* 12 (1997), pág. 17.

⁵⁵ NIELSEN, Michael A. *Neural networks and deep learning*. Vol. 2018. Determination press San Francisco, CA, 2015.

Figura 11. Propagación hacia adelante y atrás, la imagen que se encuentra en la parte superior derecha, muestra la técnica del gradiente descendiente, en la cual se muestra pequeños incrementos con el fin de encontrar el mínimo de la función. Adaptada de ⁵³



tra en la ecuación-1. El error se propaga hacia atrás, partiendo de la capa de salida hacia todas las neuronas anteriores de las capas ocultas ⁵⁶, como se describe en la ecuación-2. La importancia de este proceso consiste en que, a medida que se entrena la red, las neuronas de las capas intermedias actualicen su error de tal modo que las distintas neuronas aprenden a reconocer distintas características del espacio total de entrada, generando una variabilidad completa desde los errores totales hasta un peso específico que contribuyó a ese error.

⁵⁶ WIKIPEDIA. *Propagación hacia atrás* — *Wikipedia, La enciclopedia libre*. [Internet; descargado 19-noviembre-2020]. 2020.

$$w := w - \alpha \frac{\partial J(w)}{\partial w} \quad (1)$$

Donde:

$\frac{\partial J(w)}{\partial w}$ = Costo en función de los pesos w .

Para propagar los pesos entre capas se realiza una composición de funciones, que puede ser de la forma:

$$\frac{\partial J}{\partial w^L} = \frac{\partial J}{\partial a^L} \frac{\partial a^L}{\partial z^L} \frac{\partial z^L}{\partial w^L} \quad (2)$$

Esta composición se debe a que:

- $J(a^L)$
- $J(a^L(z^L))$
- $J(a^L(w^L a^L - 1 + b))$
- $J(a^L(w^L a^{L-1}(W^{L-1} a^{L-2} + b^{L-1}) + b^L))$

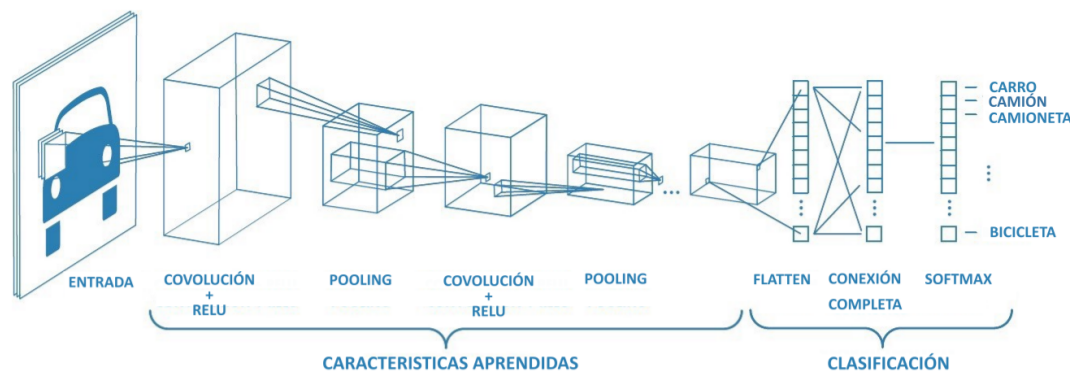
Donde: L representa la capa, a la neurona, z la salida de la neurona en la que se define la función de activación⁵⁷.

2.1.6. Redes neuronales convolucionales. Las redes neuronales convolucionales (en inglés *Convolutional Neural Networks*) (CNN) se identifican por reconocer una gran diversidad de objetos de manera autónoma. Sin embargo, se requiere de una gran cantidad de imágenes para que la red pueda captar sus características únicas.

⁵⁷ MARTINEZ, Fabio. *ai-2-UIS-student*. 2020.

Este tipo de redes identifican diferentes niveles de abstracción, por ejemplo las primeras capas logran detectar características primitivas como líneas, curvas y/o bordes, llegando a capas más profundas capaces de reconocer características más complejas; este tipo de redes neuronales son muy efectivas en aplicaciones de reconocimiento y clasificación de imágenes para visión por computadora ⁵⁸.

Figura 12. Estructura de una arquitectura de Red Neuronal Convolucional. Adaptada de ⁵⁹.



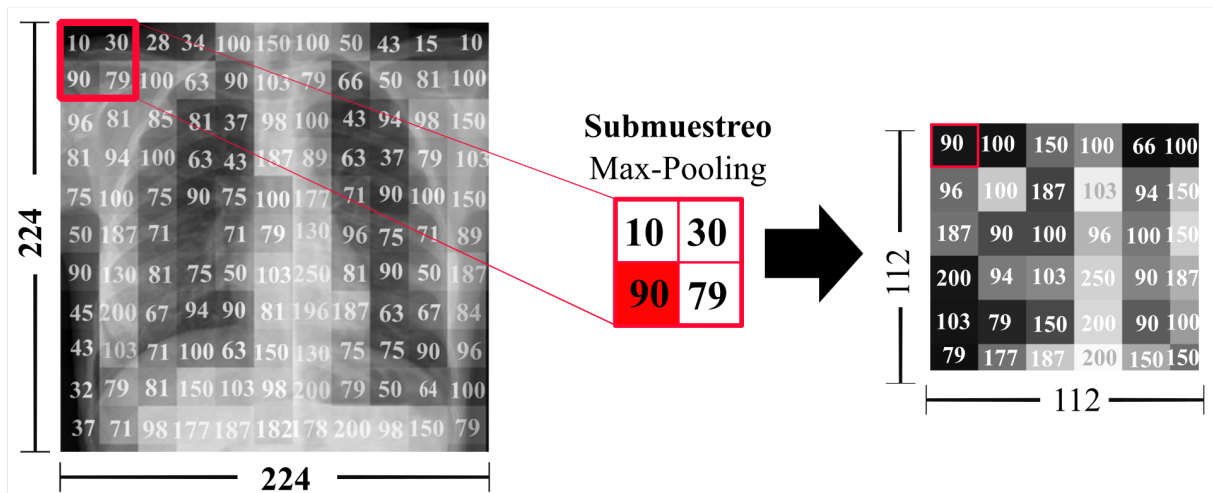
El proceso de CNN comienza con la convolución, la cual consiste en operar matemáticamente un grupo de píxeles cercanos con un kernel (por ejemplo de tamaño 3×3), que actuará como ventana deslizante y se moverá a lo alto y ancho de la imagen realizando las convoluciones entre los valores de píxeles correspondientes, adquiriendo un mapa específico de características, propio del kernel, que ayudará a la distinción de objetos ^{55,60}. Este mapa de características o matriz de salida corresponde a la siguiente capa oculta. Cada convolución genera cierto número de filtros que indica las matrices de salida.

⁵⁸ MISSINGLINK.AI. *Fully Connected Layers in Convolutional Neural Networks: The Complete Guide*. 2018.

⁶⁰ SAMIA. *Towards Machine Learning*.

Por otro lado, para reducir el número de neuronas de la próxima capa se realiza un proceso de *subsampling* que consiste en reducir las dimensiones espaciales de las imágenes filtradas predominando las características más importantes que detecto cada filtro, como apoyo en la decisión de clasificación final. Existe diversos tipos de *subsampling*, sin embargo, el más usado es el *Max-Pooling* de tamaño 2×2 , como se muestra en la figura-13.

Figura 13. Proceso de reducción de imagen con max-pooling de 2×2 , (en inglés *subsampling*).



Finalmente, el proceso de aprendizaje de una red convolucional además de una función de activación, requiere de otros hiperparámetros como el número de iteraciones de entrenamiento (épocas), esquemas de inicialización de peso, regularización, etc.

ResNet, AlexNet, VGGNet, DenseNet, son algunos modelos de redes neuronales convolucionales prometedores, reconocidos por ganar el concurso de clasificación de la base de datos ImageNet, Large Scale Visual Recognition Challenge (ILSVRC)⁶¹.

⁶¹ WIKIPEDIA CONTRIBUTORS. *ImageNet* — *Wikipedia, The Free Encyclopedia*. [Online; accessed 27-November-2020]. 2020.

Estos modelos han demostrado ser redes neuronales convolucionales efectivas para resolver problemas de clasificación de imágenes desafiantes, además se pueden reutilizar a través de la biblioteca de redes neuronales de código abierto escrita en Python, Keras.

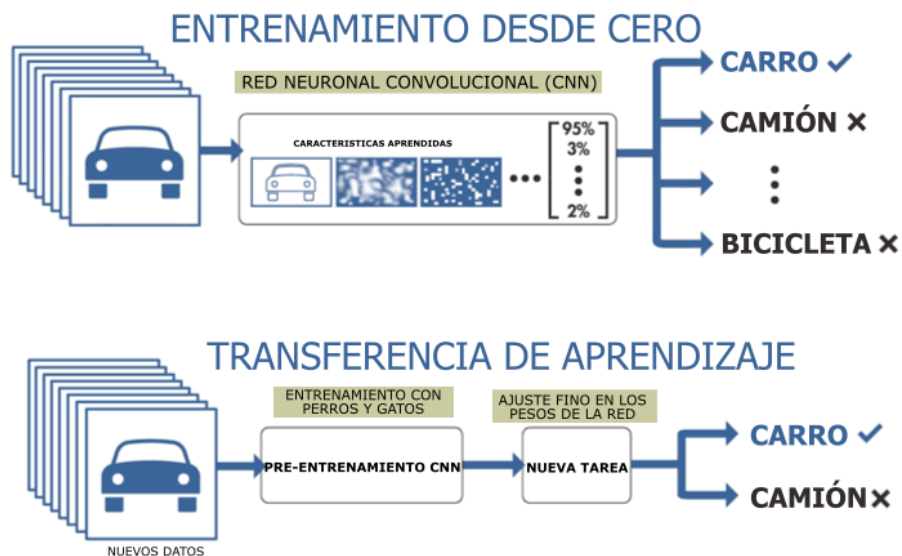
2.1.7. Transferencia de aprendizaje. El aprendizaje tradicional de las redes neuronales profundas se caracteriza por requerir una gran cantidad de muestras independientes etiquetadas y similares a los datos que más adelante se utilizarán como prueba. Sin embargo, diversos campos de acción de la inteligencia artificial cuentan con un número limitado de muestras etiquetadas, debido a altos costos y ardua recopilación de información, por lo que resulta difícil crear un conjunto de datos a gran escala y de alta calidad. Una de las áreas de aplicación en la que suele presentarse esta limitación es la medicina. Cuando una red neuronal no cuenta con una gran cantidad de datos de entrenamiento, el modelo puede que no logre generalizar las características aprendidas y presentar dificultades para identificar una nueva muestra⁶².

Para dar solución a este inconveniente, fue creada la técnica de transferencia de aprendizaje (en inglés *Transfer Learning*), la cual consiste en entrenar el algoritmo a través del uso de datos auxiliares disponibles. Es decir, no es necesario que los datos de entrenamiento y los datos de prueba pertenezcan al mismo dominio de destino. Y además, no es estrictamente necesario entrenar el modelo desde cero, sino que, existen lenguajes de programación y librerías que ya incluyen redes pre-entrenadas con un gran conjunto de imágenes naturales llamado ImageNet, que

⁶² BAZRAFKAN, Shabab, *et al.* "Deep learning for facial expression recognition: A step closer to a smartphone that knows your moods". En: *2017 IEEE International Conference on Consumer Electronics (ICCE)*. IEEE. 2017, págs. 217-220.

pueden ser cargadas y algunos de sus pesos re ajustados con una pequeña cantidad de muestras de entrenamiento que se tengan disponibles, para ajustar los pesos de la red al dominio que se desea clasificar. Esta ventaja ha incidido positivamente en el tiempo requerido para aprender completamente una nueva tarea específica. La figura-14 ilustra una comparación entre el aprendizaje tradicional y la técnica de *transfer learning*.

Figura 14. Comparación visual de un modelo preentrenado con un modelo entrenado desde cero. Adaptado de ⁶³.



El uso de la técnica de transferencia de aprendizaje no solo acelera el proceso de re-entrenamiento del modelo, sino que también, puede contribuir a obtener buenos resultados ya que se utilizan arquitecturas que han mostrado previamente un alto desempeño y que tuvieron a su disposición un conjunto de datos muy extenso para su primer entrenamiento. En⁶⁴ la autora de este proyecto presentó un enfoque ba-

⁶⁴ ESCOBAR, J. P., *et al.* "Transfer Learning with Convolutional Neural Network for Gastrointestinal Diseases Detection using Endoscopic Images". En: *2020 IEEE Colombian Conference on Applications of Computational Intelligence (IEEE ColCACI 2020)*. 2020, págs. 1-6. DOI: 10.1109/ColCACI50549.2020.9247847.

sado en el aprendizaje por transferencia para clasificar enfermedades y anomalías del tracto gastrointestinal, evidenciando que el aprendizaje automático y por transferencia son una alternativa prometedora para respaldar de forma rápida y precisa las decisiones médicas mediante el diagnóstico asistido por computadora.

2.2. ESTADO DEL ARTE

Desde el punto de vista médico, numerosos estudios han sido presentados en el estado de arte para guiar al personal del área de la salud en la detección clínica de la enfermedad de neumonía. En ⁶⁵ presentan una guía enfocada en el diagnóstico para pacientes menores de 5 años. En ⁶⁶ presentan una guía para la prevención, diagnóstico y tratamiento de neumonía en pacientes de entre 3 meses y 15 años. En ⁶⁷ evalúan las diferencias de diagnosticar la neumonía con radiografía de tórax y con tomografía computarizada. A nivel local, los estudios realizados acerca de neumonía en la Universidad Industrial de Santander son escasos. Uno de ellos ⁶⁸, describe las características en pacientes diagnosticados con neumonía con el fin de tomar medidas de control para el manejo y atención de los niños y niñas menores de 18 años que acuden al Hospital Universitario de Santander.

⁶⁵ MARTÍNEZ, J. *Factores asociados a neumonía adquirida en comunidad en infantes menores de 5 años*, Subcentro de Salud no. 1, Cuenca. 2015.

⁶⁶ MDSPD, Ecuador. "Neumonía adquirida en la comunidad en pacientes de 3 meses a 15 años". En: *Guía de Práctica Clínica* 13 (2017).

⁶⁷ SELF, Wesley H, *et al.* "High discordance of chest x-ray and computed tomography for detection of pulmonary opacities in ED patients: implications for diagnosing pneumonia". En: *The American journal of emergency medicine* 31.2 (2013), págs. 401-405.

⁶⁸ RAMIREZ RUIZ, Jose Mauricio. "Caracterización De Niños Entre 2 Meses Y 17 Años Con Diagnóstico De Neumonía Adquirida En La Comunidad Hospitalizados En El Hus." Tesis doct. Universidad Industrial de Santander, Escuela de Medicina, Departamento De . . . , 2016.

Por otra parte, uno de los temas de investigación con mayor auge en el área de la visión por computador y la inteligencia artificial, es el desarrollo y aplicación de redes neuronales convolucionales. Este es un tópico que ha tenido un crecimiento y desarrollo significativo en los últimos años, impulsado por el incremento de la capacidad computacional que hace posible el entrenamiento de arquitecturas de redes neuronales cada vez más profundas y complejas ⁷.

Particularmente, el diagnóstico de enfermedades con redes neuronales a partir de imágenes médicas y reportes biológicos es un campo de investigación activo y desafiante ⁶⁹. Específicamente, en ¹⁰ proponen una red neuronal convolucional de 121 capas para apoyar el diagnóstico de neumonía asistido por computador, mediante el etiquetado automático de radiografías de tórax. Otros estudios han comparado el rendimiento de sistemas de detección automática de neumonía bacteriana a partir de radiografías de tórax ⁷⁰. Dos de los trabajos más recientes ^{19, 15} proponen enfoques con aprendizaje profundo para la detección de neumonía en imágenes de rayos-X utilizando el concepto de transferencia de aprendizaje, en ¹⁹ los autores registran una exactitud del 98 % para la detección de neumonía clasificado en normal y neumonía. En ¹⁵ presenta un enfoque similar alcanzando el 96 % de exactitud.

En este sentido, a nivel local, la Universidad Industrial de Santander reporta múltiples estudios a partir de redes neuronales e inteligencia artificial, especialmente

⁶⁹ LEE, Honglak , *et al.* "Unsupervised feature learning for audio classification using convolutional deep belief networks". En: *Advances in neural information processing systems*. 2009, págs. 1096-1104.

⁷⁰ FISZMAN, Marcelo , *et al.* "Automatic detection of acute bacterial pneumonia from chest X-ray reports". En: *Journal of the American Medical Informatics Association* 7.6 (2000), págs. 593-604.

desde la Escuela de Ingeniería de Sistemas e Informática. Por ejemplo, en ⁷¹ implementan y evalúan métodos con redes neuronales convolucionales para la clasificación de la enfermedad de Alzheimer. En ⁷² predicen el rendimiento agrícola de cultivos de cacao, detección de daños en puentes ⁷³, entre otros. Sin embargo, ninguno de ellos ha estado orientado a la detección automática de neumonía en imágenes.

Con base en lo expuesto anteriormente, se considera apropiado para la formación de la autora de este trabajo como Ingeniera de Sistemas, abordar el problema de procesamiento y clasificación de imágenes de rayos-X de tórax para la detección de neumonía haciendo uso de redes neuronales convolucionales.

2.3. MARCO TECNOLÓGICO

2.3.1. Software. En términos de Software, el presente trabajo emplea el lenguaje de programación Python en Google Colab. La herramienta implementa GPU, proporcionada de manera gratuita. Además, se hizo uso de herramientas básicas de ofimática para el desarrollo de los reportes y presentación final de resultados, tales como licencia de Windows, Microsoft word, y Microsoft Power Point. Todas ellas son patrocinadas por la Universidad Industrial de Santander.

⁷¹ ORTIZ SUAREZ, Juan Manuel. "Evaluacion De Metodos Con Redes Neuronales Convolucionales Para La Clasificacion De La Enfermedad De Alzheimer En Imagenes De Resonancia Magnetica Cerebral." Tesis doct. Universidad Industrial de Santander, Escuela De Ing. De Sistemas, 2016.

⁷² GOMEZ, Maria y JAIMES, Estefania. "OPTIMIZACION PARAMETROS REDES NEURONALES ARTIFICIALES PREDICCION RENDIMIENTO AGRICOLA CULTIVOS CACAO MEDIANTE METODOS SELECCION INGENIERIA CARACTERISTICAS". Tesis doct. Universidad Industrial de Santander, Escuela De Ing. De Sistemas, 2019.

⁷³ MARTINEZ, Hani y EDUARDO, Rafael. "Deteccion De Danos En Puentes De Tipologia Viga Y Losa Simplemente Apoyada Usando Redes Neuronales Artificiales." Tesis doct. Universidad Industrial de Santander, Escuela De Ing. Civil, 2012.

2.3.2. Hardware. La autora del presente trabajo es integrante activa del Grupo de Investigación en diseño de algoritmos y procesamiento de datos multidimensionales (HDSP), adscrito a la Escuela de Ingeniería de Sistemas e Informática de la UIS. Por lo tanto, dispuso de los recursos de trabajo con que cuenta el grupo de investigación para llevar a cabo el desarrollo del proyecto. Específicamente, el equipo mediante el cual se desarrolló el proyecto es:

- Computador de escritorio HP 600 G1 SFF CORE 17-4790 QUAD Core 3.60 Ghz. disco duro 1TB 72000 RPM Memoria RAM 16 GB DDR3 1600 Mhz, Sistema operativo Windows 7 Profesional, 64 con Windows Pro Licence 3-3-3.

3. METODOLOGÍA

En este estudio se busca desarrollar mediante transferencia de aprendizaje un algoritmo para procesar imágenes médicas con el fin de proporcionar un acceso de diagnóstico oportuno para la detección de neumonía a partir de imágenes de rayos-X de tórax. El algoritmo se desarrolla en el lenguaje de programación Python y se ejecutó en Google Colab⁷⁴.

Objetivo específico No. 1. "Seleccionar una base de datos de radiografías de tórax en humanos con y sin neumonía, etiquetada por personal médico experto, disponible en internet y avalada para fines de investigación."

3.1. BASE DE DATOS

Para dar cumplimiento al objetivo específico No.1, la primera etapa de la metodología consiste en realizar una búsqueda exhaustiva de bases de datos de imágenes de rayos-X, de dominio público, y aceptados en revistas científicas que avalen su uso para fines de investigación. Además, se espera que el conjunto de datos este constituido por una cantidad de imágenes suficiente que permita el desarrollo de un sistema de clasificación por aprendizaje profundo, aún teniendo en cuenta la técnica de transferencia de aprendizaje. La base de datos seleccionada en esta fase de la metodología se expone en el primer numeral de la sección No. 4. Resultados.

Objetivo específico No. 2. "Diseñar una estrategia de procesamiento y clasi-

⁷⁴ BISONG, Ekaba. "Google Colaboratory". En: *Building Machine Learning and Deep Learning Models on Google Cloud Platform*. Springer, 2019, págs. 59-64.

ficación de las imágenes de rayos-X con y sin neumonía a partir de redes neuronales convolucionales.”

De acuerdo con lo establecido en el objetivo específico No. 2, se exploraron diversas técnicas de procesamiento, a fin de elegir las más apropiadas para conformar una estrategia que permita la clasificación de la base de datos seleccionada en la primera fase de la metodología. Las estrategias de procesamiento de imágenes exploradas se exponen a continuación:

3.2. PROCESAMIENTO DE IMÁGENES

3.2.1. Aumento de datos. El aumento de datos consiste en realizar transformaciones a las imágenes originales que conforman un conjunto de datos sin realizar modificaciones en las etiquetas. El objetivo es generar nuevas imágenes que guarden una relación válida con las originales con el fin de tener mayor cantidad de muestras para hacer las tareas de entrenamiento y clasificación. Algunas de las técnicas de procesamiento más usadas para generar nuevas imágenes en el área de aprendizaje profundo son las siguientes:

- Rotación (*rotation_range*): Consiste en girar la imagen en el sentido de las agujas del reloj, en un ángulo entre 0 y 360 grados.
- Escalado (*rescale*): Factor de cambio de escala, radica en ampliar o reducir el tamaño de la imagen mediante la multiplicación del valor proporcionado.
- Traslación (*Flip*): Voltea las entradas horizontal (*horizontal_flip*) o verticalmente (*vertical_flip*).
- Enfocar (*zoom_range*): Se emplea para acercar o alejar las imágenes.

- Cambio de ancho (*width_shift_range*): Indica el desplazamiento horizontal de la imagen, hacia la izquierda o hacia la derecha de forma aleatoria.

En la página web de la API de Keras⁷⁵ se pueden consultar todas las transformaciones implícitas en la librería Keras de Python, las cuales pueden ser implementadas a través de la biblioteca de Keras *ImageDataGenerator*.

Por otra parte, el origen de uso de esta técnica radica en la hipótesis de que a mayor cantidad de muestras, se espera que mejore el aprendizaje del modelo ⁷⁶. Sin embargo, la implementación de técnicas de aumento de datos debe elegirse cuidadosamente, valiéndose del contexto y dominio del problema, ya que el uso de algunas de las técnicas de transformación implican la creación de nuevas imágenes alejadas de la realidad, lo cual no es ampliamente aceptado en algunas áreas de estudio ^{77,78}. Específicamente, varios autores reconocidos del área de investigación en imágenes médicas, aseguran que no siempre es recomendable manipular o modificar los datos de entrada con este tipo de técnicas, ya que se puede incurrir en la generación de imágenes humanas alejadas de la realidad, que alterarían negativamente el aprendizaje de un modelo, el cual ya no correspondería legítimamente a imágenes reales humanas.

⁷⁵ ABADI, Martín, *et al.* *TensorFlow: Large-Scale Machine Learning on Heterogeneous Systems*. Software available from tensorflow.org. 2015.

⁷⁶ WONG, S. C., *et al.* "Understanding Data Augmentation for Classification: When to Warp?" En: *2016 International Conference on Digital Image Computing: Techniques and Applications (DICTA)*. 2016, págs. 1-6. DOI: 10.1109/DICTA.2016.7797091.

⁷⁷ WONG, Sebastien C, *et al.* "Understanding data augmentation for classification: when to warp?" En: *2016 international conference on digital image computing: techniques and applications (DICTA)*. IEEE. 2016, págs. 1-6.

⁷⁸ VAN DYK, David A y MENG, Xiao-Li. "The art of data augmentation". En: *Journal of Computational and Graphical Statistics* 10.1 (2001), págs. 1-50.

3.2.2. Ecuación de histogramas. El algoritmo de ecualización consiste en utilizar la función de transformación *cv2.equalizeHist()*, para definir una misma condición de iluminación y aumentar el contraste para todas las imágenes, es decir el número de píxeles por cada nivel de intensidad sea extendido en todo el rango $[0 - 255]$. Esta función recibe como entrada una imagen a escala de grises, a través de la librería de OpenCv en Python. La técnica consiste en leer las imágenes, seguidamente para asegurar que las imágenes estén trabajando con un canal espectral, se convierte cada una de las imágenes originales a escala de grises, luego se implementa el algoritmo mediante la función de OpenCV *equalizeHist*, finalmente se guardan las imágenes.

Además de explorar técnicas de procesamiento de imágenes, en esta fase de la metodología se indaga acerca de las redes neuronales convolucionales que permitirían la solución del problema de clasificación binaria abordado en el presente proyecto. En la siguiente fase (3.3) se expone el funcionamiento de las redes neuronales pre-seleccionadas en esta segunda fase (3.2).

3.2.3. Redimensión de imágenes. Cada imagen del conjunto de datos tiene un tamaño espacial distinto. La operación de convolución implica un volumen de entrada fijo y un tamaño de filtro fijo, por esta razón, es necesario que las imágenes sean del mismo tamaño, antes de alimentar la red. Teniendo en cuenta la resolución de la capa de entrada de la red neuronal VGA, se decide dimensionar todas las imágenes a 224×224 píxeles espaciales, estas dimensiones fueron establecidas para todas las arquitecturas. La estrategia de procesamiento y clasificación de las imágenes desarrollada es expuesta en la sección 4.2. Resultados.

Objetivo específico No. 3. “Implementar un algoritmo para ejecutar la estrate-

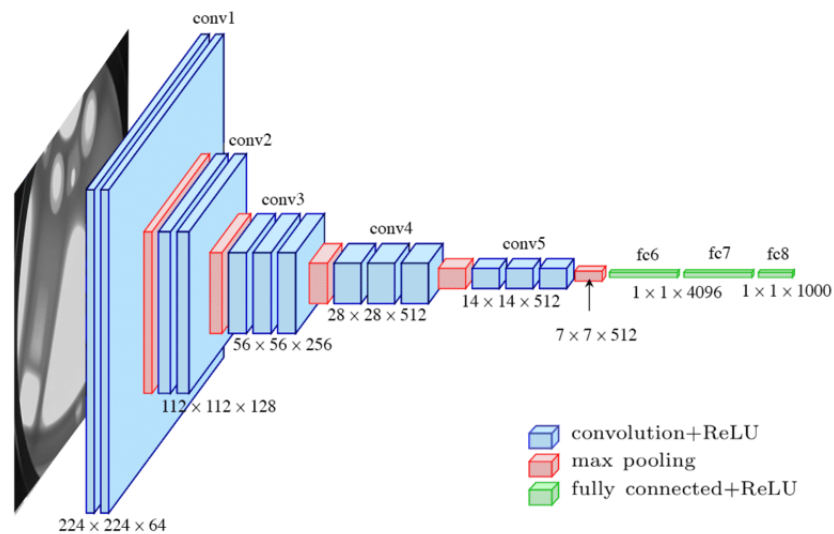
gia diseñada, incorporando la técnica de transferencia de aprendizaje.”

Para desarrollar el objetivo específico No. 3, se describe a continuación, cuatro arquitecturas de CNNs del estado de arte que fueron seleccionadas durante el desarrollo de la fase 2 (3.2). Estas cuatro arquitecturas de CNNs serán implementadas en Python durante el desarrollo de esta tercera fase de la metodología. El propósito es comparar el rendimiento de la clasificación de la base de datos, al realizar la implementación de cada una de las cuatro arquitecturas por separado. De esta manera, se concluirá cuál CNN es la más apropiada para abordar el problema de este proyecto.

3.3. MODELO DE APRENDIZAJE PROFUNDO PARA LA CLASIFICACIÓN DE NEUMONÍA

3.3.1. Arquitectura VGG-16. La arquitectura VGG16 fue la ganadora del desafío de reconocimiento visual a gran escala de la base de datos ImageNet en el año 2014 (ILSVRC).

Figura 15. Arquitectura convolucional del modelo VGG-16. Tomado de⁷⁹.



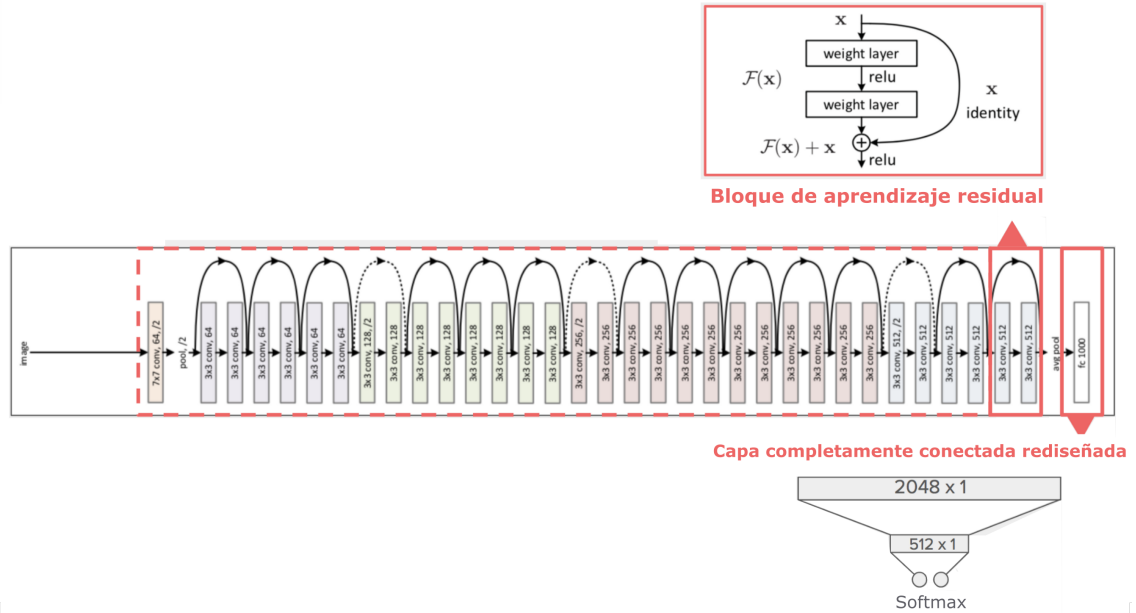
La red VGG16, desarrollada por Simonyan y Zisserman ⁸⁰, está conformada por 16 capas convolucionales. La figura-15, muestra una serie de cubos de color azul, los cuales representan las capas de convolución, dichos filtros de convolución son pequeños, específicamente, de tamaño (3×3) . En total se ilustran 13 cubos azules, en representación de las 13 capas de convolución, y adicionalmente se observa de color rojo, 5 capas de agrupación máxima, de tamaño (2×2) ; y 3 cubos verdes que describen las 3 capas completamente conectadas, para un total de 16 capas convolucionales que componen a esta red. El tamaño de entrada predeterminado para este modelo es 224×224 y todas las capas están completamente conectadas, contiene 138, 357, 544 parámetros entrenables, las capas ocultas tienen ReLU como función de activación.

3.3.2. Arquitectura ResNet-50. La idea central de las arquitecturas ResNet es aprender la función residual, el número de parámetros de ResNets es sustancialmente mayor. Este modelo es conformado por una arquitectura convolucional de 50 capas de profundidad, opcionalmente, carga pesos previamente entrenados en ImageNet⁸¹ además, fue la primera versión de redes muy profundas, resolviendo el problema de desvanecimiento del gradiente durante el entrenamiento, a través de conexiones residuales que transfiere información intercalada entre capas.

⁸⁰ SIMONYAN, Karen y ZISSERMAN, Andrew. "Very deep convolutional networks for large-scale image recognition". En: *arXiv preprint arXiv:1409.1556* (2014).

⁸¹ HE, Kaiming, *et al.* "Deep residual learning for image recognition". En: *Proceedings of the IEEE conference on computer vision and pattern recognition*. 2016, págs. 770-778.

Figura 16. Arquitectura ResNet con 50 capas.



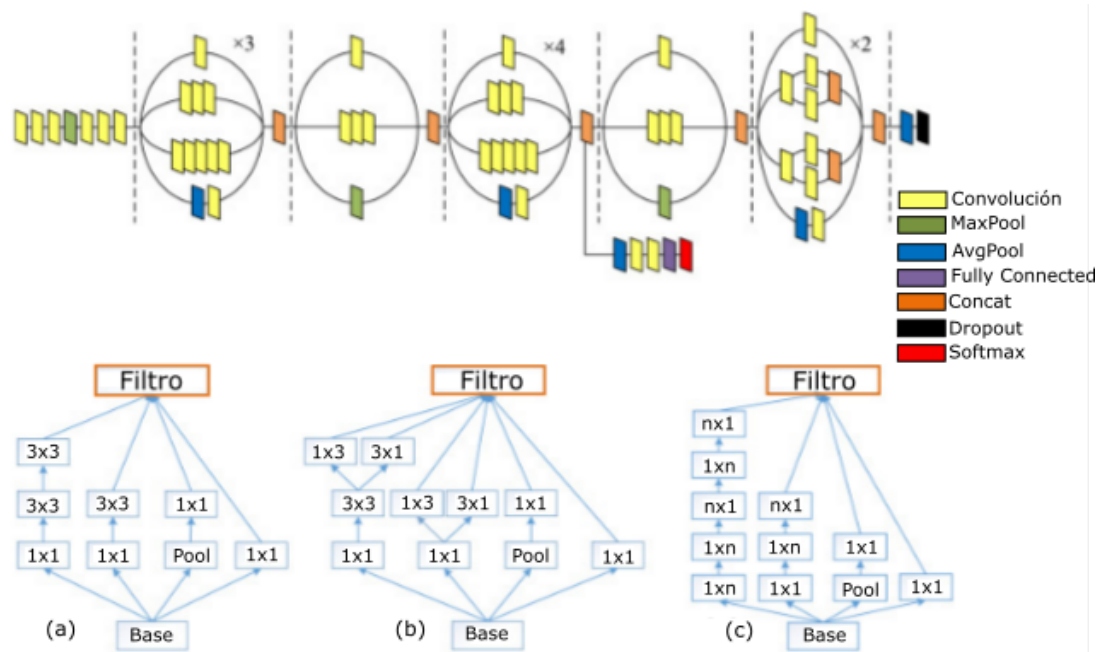
3.3.3. Arquitectura Inception V3. La arquitectura Inception-V3 de Google Net es una CNN propuesta por Google en 2015, diseñada específicamente para problemas de reconocimiento y clasificación de imágenes.

La primera versión de inception fue creada en el año 2014, y fue nombrada, Inception-V1. La figura-17 ilustra el diagrama estructural del modelo de Inception-v3. Esta arquitectura es una estructura de red disponible en Keras, que está previamente entrenada con el conjunto de datos de ImageNet⁸³, donde puede reconocer más de 1000 categorías de datos de imágenes diferentes. El tamaño de entrada por defecto es de $[299 \times 299 \times 3]$, este modelo utiliza un método de división del kernel de convolución con el fin de reducir el número de parámetros⁸⁴.

⁸³ DENG, Jia, *et al.* "Imagenet: A large-scale hierarchical image database". En: *2009 IEEE conference on computer vision and pattern recognition*. IEEE. 2009, págs. 248-255.

⁸⁴ WANG, Cheng, *et al.* "Pulmonary image classification based on inception-v3 transfer learning model". En: *IEEE Access* 7 (2019), págs. 146533-146541.

Figura 17. La imagen representa la arquitectura convolucional Inception version 3, se presenta los detalles del bloque. La figura-17 (a) presenta un módulo para capas iniciales con factorización simétrica, la figura-17 (b) representa un módulo inception para capas finales para representaciones de alta dimensionalidad, finalmente la figura17 (c) describe un módulo para capas intermedias con factorización asimétrica. Adaptada de ⁸²

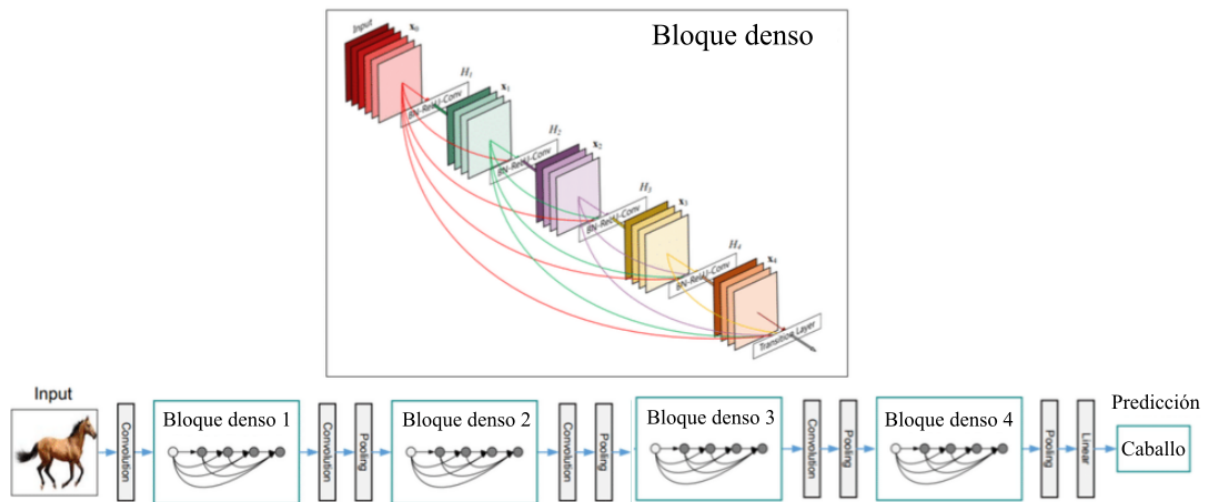


3.3.4. Arquitectura DenseNet-169. Esta arquitectura ganó el primer lugar en el reto de clasificación del conjunto de datos ImageNet, realizado en el marco de la conferencia de visión por computador y reconocimiento de patrones CVPR 2017⁸⁵. Hasta el momento la arquitectura tiene tres versiones diferentes: DenseNet-121, DenseNet-169 y DenseNet-201, cuyo nombre varía de acuerdo con el número de capas que contiene. La red DenseNet-169 solventa el problema del desvanecimiento de gradiente debido a su patrón de conectividad denso. Las redes DenseNet presentan un efecto de regularización particular, que les permite reducir el sobreajuste con conjuntos de entrenamiento pequeños. Además, se caracterizan por la

⁸⁵ HUANG, Gao, *et al.* "Densely connected convolutional networks". En: *Proceedings of the IEEE conference on computer vision and pattern recognition*. 2017, págs. 4700-4708.

concatenación de mapas de características de todos los bloques convolucionales precedentes, lo que constituye una gran diferencia entre las arquitecturas convolucionales DenseNet y ResNet.

Figura 18. La imagen representa la arquitectura DenseNet de 4 capas, con 4 bloques. Tomado de⁸⁶.



En términos de configuración, la red DenseNet-169 tiene un tamaño de entrada por defecto de 112×112 , cuatro bloques densos cada uno con el mismo número de capas y funciones de activación lineales. Antes de que las muestras ingresen al primer bloque denso hay una convolución de tamaño 7×7 , y un agrupamiento promedio (en inglés *global average pooling*) de 3×3 . Luego, entre bloques densos hay una convolución 1×1 y un *global average pooling* de 2×2 . Después del último bloque denso hay un *global average pooling*, seguido de la capa completamente conectada para realizar la clasificación con función de activación softmax.

En comparación con las redes Inception y ResNet, que también concatenan características de diferentes capas, las arquitecturas DenseNet reducen sustancialmente el número de parámetros, ya que no hay necesidad de volver a aprender mapas de características redundantes. Por otra parte, debido a que el valor óptimo de un

hiperparámetro para cierta tarea no se puede conocer a priori y es muy dependiente del problema en cuestión, se procedió a encontrar los valores más apropiados para la red de manera heurística. Los valores se describen en la tabla-3 sección 4.4.

A continuación, se presenta una breve descripción de los hiperparámetros modificados en cada una de las cuatro arquitecturas:

Batch-Size: Determina el número de datos que usaremos para cada actualización, consiste en determinar cuántas muestras de la base de datos de entrenamiento se van a tomar en cada iteración, actualiza los hiperparámetros de la red y obtiene el gradiente en dicho análisis.

Dropout: Se basa en eliminar estocásticamente neuronas durante el entrenamiento para evitar la coadaptación de los detectores de características. Su objetivo es prevenir el overfitting en la red.

Tasa de aprendizaje: Los algoritmos de gradiente descendiente, multiplican la magnitud del gradiente por un escalar conocido como *learning rate* (también denominado, a veces, *step size*) para determinar el siguiente punto. En general, si este valor es demasiado grande, se harán avances enormes y procederá un rápido proceso de aprendizaje sin embargo, es posible que se salte el mínimo y dificulte el proceso de aprendizaje. Por el contrario, si la tasa de aprendizaje es pequeña, se harán avances constantes y pequeños, teniéndose una mejor oportunidad de llegar a un mínimo local, pero esto puede provocar que el proceso de aprendizaje sea muy lento.

Optimizador: La función del optimizador como su palabra lo indica, es optimizar los valores de los hiperparámetros, para reducir el error cometido por la red mediante la actualización de los valores, con base a la tasa de aprendizaje. El proceso mediante el cual se hace esto, se conoce como "*backpropagation*".

Epochs: Este hiperparámetro indica el número de veces que se ejecutará el algoritmo de *forward-propagation* y *back-propagation*. Este valor permite continuar iterando en los anteriores pasos. A mayor número de épocas mayor costo computacional.

Objetivo específico No. 4. “Evaluar el rendimiento del algoritmo de procesamiento y clasificación de imágenes de rayos-X para el diagnóstico de neumonía en términos de tiempo y exactitud.”

Con el propósito de evaluar el rendimiento del algoritmo propuesto tal como se plantea en el objetivo específico No. 4, durante el desarrollo de la cuarta fase del proyecto se implementarán las métricas descritas en el párrafo siguiente. Adicionalmente, se registrará la RAM consumida (GB), uso en disco (GB), rendimiento a través del consumo de memoria en unidades de procesamiento de gráficos (en inglés *Graphic processing unit (GPU)*) (GB), y el tiempo (s).

3.4. MÉTRICAS DE VALIDACIÓN

Una vez que se evalúa el método de clasificación con los datos de prueba o *test*, el modelo de clasificación asigna a cada imagen una etiqueta, ya sea Normal o Neumonía, según la clase con la cual el modelo identifique a la imagen. Dependiendo de si la etiqueta es correcta o incorrecta, se puede calcular para la totalidad de datos de prueba o *test* los siguientes valores: Verdadero positivo (en inglés, *True Positive -TP*); TP representa el número total de elementos de una clase X predicha correctamente; Verdadero negativo (en inglés, *True Negative -TN*), corresponde al número total de elementos de otra clase predichos correctamente; falso positivo (en inglés, *False Positive -FP*), el cual corresponde al número total de elementos predichos en una clase X pero pertenecientes a otra clase; falso negativo (en inglés, *False Nega-*

tive -FN), representa el total de elementos erróneos de una clase X predichos con cualquier otra clase⁸⁷.

Teniendo en cuenta las etiquetas mencionadas, se puede calcular las métricas descritas a continuación, para evaluar el rendimiento del algoritmo.

3.4.1. Exactitud (Accuracy). La ecuación-3 se refiere a una medida de sesgo estadístico, error absoluto o relativo de una cantidad aproximada, que se utiliza específicamente, para describir la proximidad de una medida al valor real⁸⁸. La exactitud en específico, se refiere a las predicciones correctas que ha hecho el modelo y el total de predicciones.

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \quad (3)$$

3.4.2. Precisión (PRE). La ecuación-4 calcula la precisión de las predicciones con respecto a las etiquetas. La métrica utiliza TP y FP para realizar el cálculo⁸⁹. Este valor se devuelve finalmente como precisión, el cual consiste en dividir TP por la suma de TP y FP. Si el valor de la precisión es igual a uno, se infiere que todas

⁸⁷ GOMEZ, Natalia Andrea; SANCHEZ, Karen y ARGUELLO, Henry. "Non-Destructive Method for Classification of Cocoa Beans from Spectral Information". En: *2019 XXII Symposium on Image, Signal Processing and Artificial Vision (STSIVA)*. IEEE. 2019, págs. 1-5.

⁸⁸ HUANG, Jin y LING, Charles X. "Using AUC and accuracy in evaluating learning algorithms". En: *IEEE Transactions on knowledge and Data Engineering* 17.3 (2005), págs. 299-310.

⁸⁹ ABADI, Martín, *et al.* *TensorFlow: Large-Scale Machine Learning on Heterogeneous Systems*. Software available from tensorflow.org. 2015.

las precisiones con respecto a las etiquetas son correctas.

$$Pre = \frac{TP}{TP + FP} \quad (4)$$

3.4.3. Exhaustividad (*Recall*). También conocida como sensibilidad, esa métrica compara las etiquetas identificadas como verdaderos FN sobre el total de verdaderos positivos TP. En general, es la proporción de casos positivos que fueron correctamente identificadas por el algoritmo, y se calcula según la ecuación-5.

$$Recall = \frac{TP}{TP + FN} \quad (5)$$

3.4.4. Matriz de confusión. Una de las métricas más intuitivas para evaluar el rendimiento de un algoritmo de clasificación es la matriz de confusión⁹⁰.

La matriz cuenta los tipos de predicciones correctas e incorrectas que hace el clasificador para cada una de las clases. La dimensión es $n \times m$, con $n = m$, donde n representa la clase observada y m , la clase predicha. Uno de los beneficios de las matrices de confusión es la fácil visualización de resultados, debido a que muestra explícitamente cuando una clase se confunde con otra. La matriz mostrada a continuación ilustra un problema específico de clasificación binario, perpetuando que la matriz de confusión se basa en lo aprendido en instancias anteriores.

TN	FP
FN	TP

Objetivo específico No. 5. “Comparar los resultados obtenidos con otros métodos similares del estado del arte.”

⁹⁰ GOWER, John C. “A general coefficient of similarity and some of its properties”. En: *Biometrics* (1971), págs. 857-871.

Durante el desarrollo de la quinta fase del proyecto, se realizará una revisión exhaustiva de trabajos relacionados con el enfoque propuesto. Teniendo en cuenta la base de datos seleccionada en la primera fase, se dará prioridad en este ítem a aquellos trabajos que evalúen su rendimiento sobre ese mismo conjunto de datos. Se analizará la precisión, así como la optimización de recursos y costos computacionales. La tabulación de esta comparación con los resultados de otros autores se exponen en la sección No. 4, numeral quinto (4.5)

4. RESULTADOS

Objetivo específico No. 1. “Seleccionar una base de datos de radiografías de tórax en humanos con y sin neumonía, etiquetada por personal médico experto, disponible en internet y avalada para fines de investigación.”

4.1. Selección de base de datos.

Como resultado del primer objetivo específico se seleccionó el conjunto de datos **Chest X-Ray Images (Pneumonia)**, el cual contiene 5.856 imágenes de rayos-X, adquiridas en el Centro Médico de Niños y Mujeres de Guangzhou, en China⁹¹. La base de datos en mención es de acceso público gratuito y las imágenes se encuentran divididas o etiquetadas en dos categorías: Normal o Neumonía. Dentro de la categoría de imágenes con *Neumonía* existen casos de neumonía viral y bacteriana. Sin embargo, esta subdivisión (viral y bacteriana) no será tomada en cuenta en el presente proyecto.

Por otra parte, el conjunto de datos **Chest X-Ray Images (Pneumonia)** se encuentra distribuido por defecto en 3 subconjuntos, entrenamiento (Train), validación (Value) y prueba (Test), así como se muestra en la tabla-2. La figura-19, por su parte, presenta una captura de pantalla del sitio web en el cuál se encuentra disponible el conjunto de datos para su libre descarga.

⁹¹ MOONEY, Paul. *Chest X-Ray Images (Pneumonia)*. 2018.

Figura 19. Captura de pantalla de la dirección web en la cual se encuentra disponible el conjunto de datos Chest X-Ray Images (*Pneumonia*). Disponible en línea: <https://www.kaggle.com/paultimothymooney/chest-xray-pneumonia>

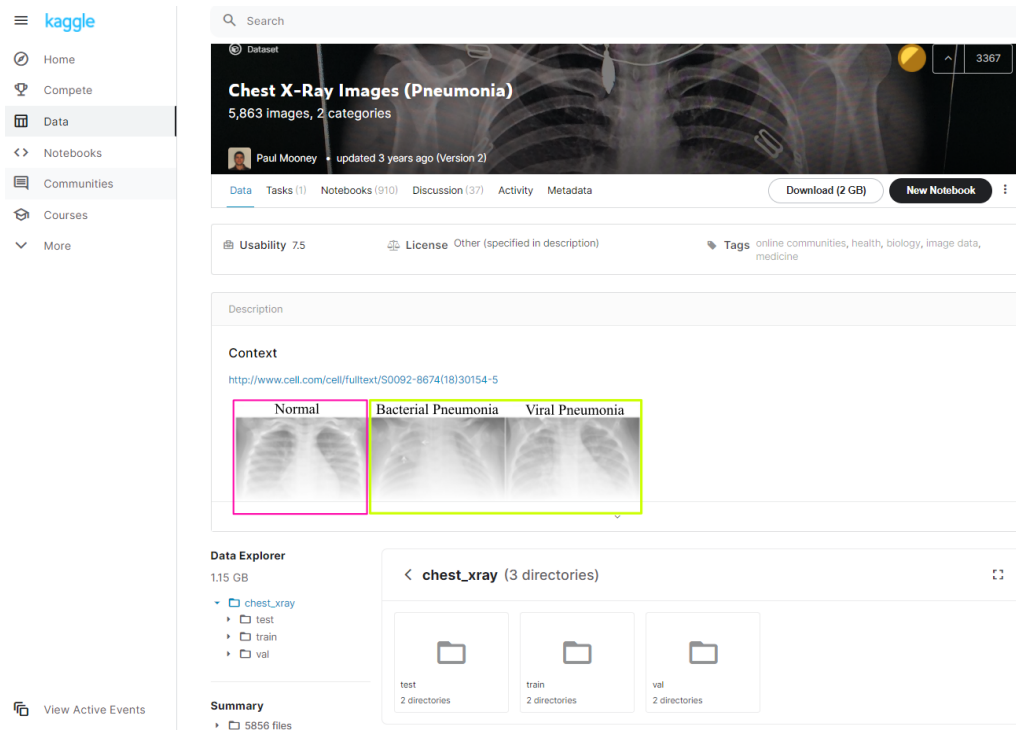


Tabla 2. Cantidad de imágenes por categoría y destino del conjunto de datos Chest X-Ray Images (*Pneumonia*) de Guangzhou, China⁹².

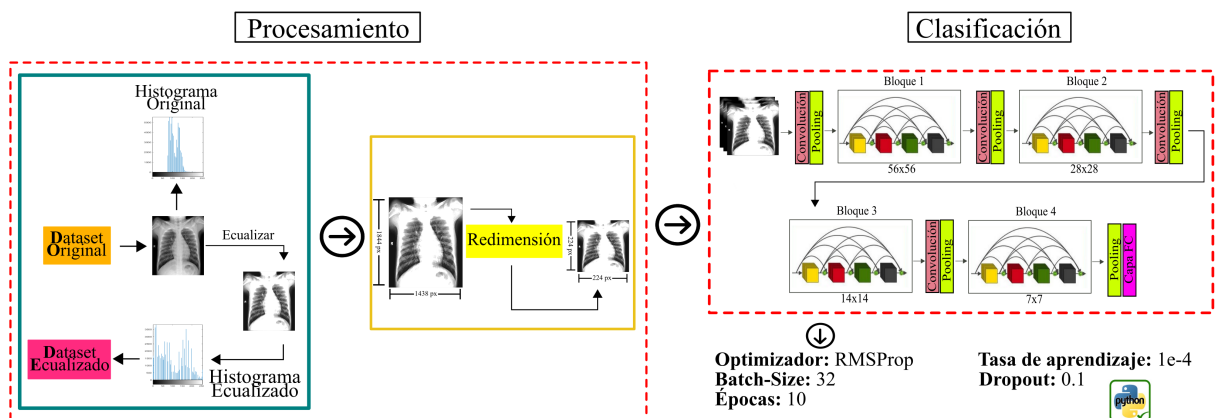
	%	No. Imágenes Normal	No. Imágenes Neumonía
Train	89	1341	3871
Value	1	8	8
Test	10	234	390
Total	100	1583	4269

Objetivo específico No. 2. “Diseñar una estrategia de procesamiento y clasificación de las imágenes de rayos-X con y sin neumonía a partir de redes neuronales convolucionales.”

4.2. Estrategia de Procesamiento y Clasificación

Como resultado de la segunda etapa de la metodología, el presente proyecto propone la siguiente metodología: Implementación de las técnicas de **ecualización** y **redimensión del conjunto de datos**, como etapas de procesamiento para el dataset mencionado en 4.1. Seguido de su clasificación binaria mediante la red neuronal convolucional **DenseNet-169**, usando el Optimizador RMSProp con tasa de aprendizaje $1 \text{ e-}4$, batch size de 32, y un Dropout de 0.1. La descripción técnica de la Ecualización es expuesta en la sección 3.2.2. Además, los detalles de los experimentos de las 4 redes neuronales a partir de los cuales se seleccionó la arquitectura densenet-169 son expuestos en la sección 4.5. Las simulaciones que permitieron definir la estrategia de clasificación en mención son detalladas en la sección 4.4. Por otra parte, es de mencionar que la red DenseNet-169 fue seleccionada luego de realizar comparaciones entre las cuatro arquitecturas descritas en la sección 3.3. Además, la versión a implementar de cada una de las arquitecturas (VGG, ResNet, Inception, y DenseNet) se definió luego de simulaciones preliminares que mostraron mejor rendimiento en términos de exactitud de clasificación.

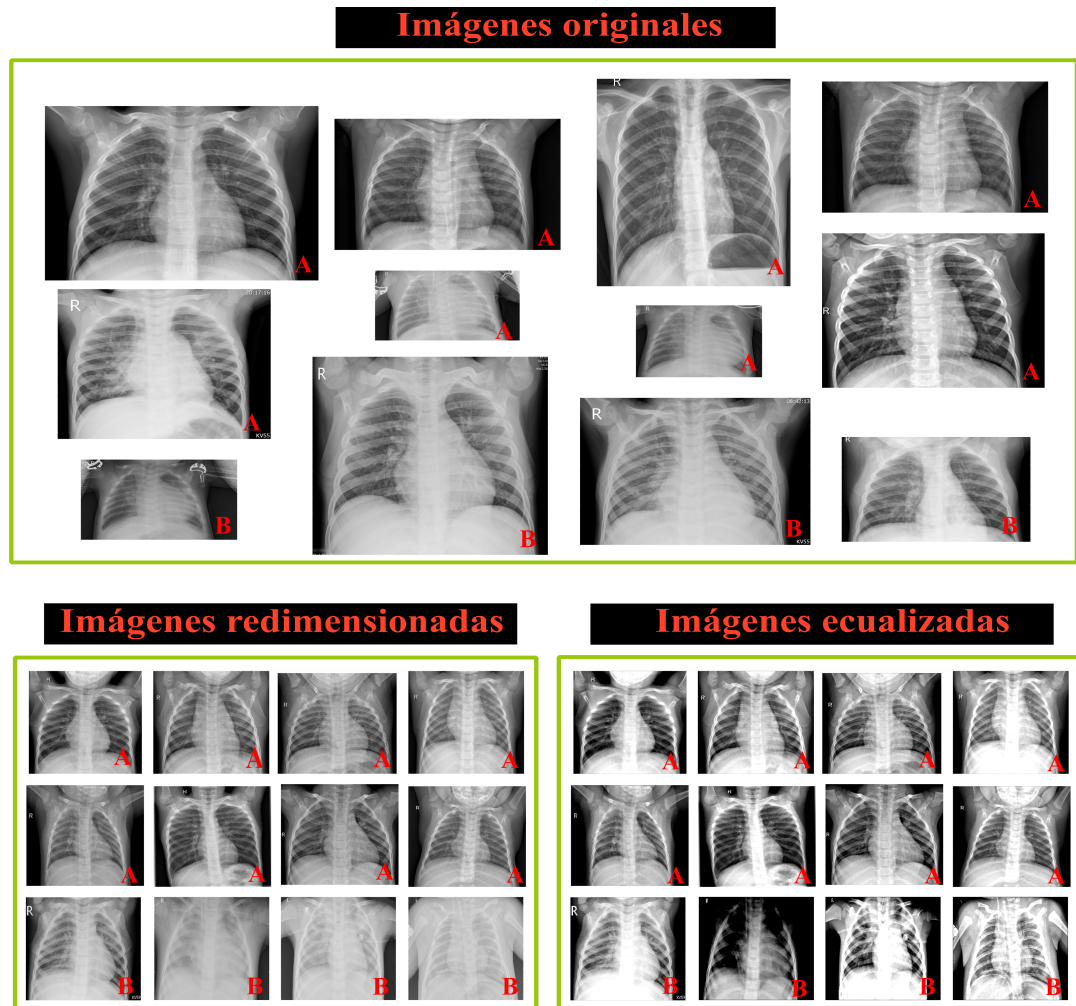
Figura 20. Estrategia propuesta. A la izquierda, en Procesamiento, se ilustran las técnicas de ecualización y redimensión, respectivamente. A la derecha, la arquitectura de clasificación DenseNet-169 empleada, y los hiperparámetros definidos.



4.3. Pre-procesamiento.

La figura-21 ilustra una colección de imágenes aleatorias del conjunto de datos Chest X-Ray Images (*Pneumonia*) de Guangzhou, China. Muestras originales se ubican a la superior, en la inferior derecha se ubican las imágenes redimensionadas a 224×224 y su resultado después de realizar la ecualización de histogramas a la izquierda.

Figura 21. Muestra un conjunto de imágenes tomadas del conjunto de datos Chest X-Ray Images (*Pneumonia*) de Guangzhou, China⁹³. Dataset original (superior), procesamiento (Inferior). (A) muestra casos normales, (B) muestras casos de neumonía.



Objetivo específico No. 3. “Implementar un algoritmo para ejecutar la estrategia diseñada, incorporando la técnica de transferencia de aprendizaje.”

4.4. Implementación del algoritmo

Una vez revisada la literatura y preseleccionadas las técnicas de procesamiento y clasificación de imágenes a implementar, se desarrollaron en lenguaje de programación Python las siguientes simulaciones.

Implementación de las dos técnicas de procesamiento: Ecuilización y Redimensionamiento; e implementación de cuatro redes neuronales convolucionales por separado: VGG-16, ResNet-50, Inception-V3 y DenseNet-169. En la sección 4.5 se describe el comportamiento y rendimiento de cada una. Posteriormente, se procedió al ajuste de los hiperparámetros, así: se varió la tasa de aprendizaje entre 0,001 y 0,0001; el valor de lote o *batch-size* entre 10 y 35; un dropout antes de la capa completamente conectada entre 0 y 0,5 como solución al problema de sobre-entrenamiento. Además, se mantuvo un valor constante respecto al número de épocas (10). Se implementaron los optimizadores Adam y RMSProp en cada red neuronal para encontrar el más adecuado para cada red al desarrollar la tarea de clasificación binaria presentada en este documento.

Luego de ejecutar las simulaciones con todas las variaciones mencionadas en el párrafo inmediatamente anterior, los diferentes hiperparámetros definidos para cada uno de los modelos CNN (VGG-16, ResNet-50, Inception-V3 y DenseNet-169) son mostrados en la tabla-3.

Una vez encontrados los hiperparámetros más apropiados para cada red neuronal, en términos de exactitud, se procedió a realizar 8 ejecuciones consecutivas y

computar su promedio. Los resultados alcanzados son expuestos en las Tablas 4 y 5 de la sección 4.5.

Tabla 3. Hiperparámetros definidos para el entrenamiento de cada modelo.

Modelo CNN	Optimizador	Batch Size	Tasa de Aprendizaje	Dropout	No. de Épocas
VGG-16	Adam	20	1e-3	0.1	10
ResNet-50	Adam	16	1e-3	0.3	10
Inception-V3	RMSProp	25	1e-4	0.1	10
DenseNet-169	RMSProp	32	1e-4	0.1	10

Objetivo específico No. 4. "Evaluar el rendimiento del algoritmo de procesamiento y clasificación de imágenes de rayos-X para el diagnóstico de neumonía en términos de tiempo y exactitud."

4.5. Evaluación del rendimiento del algoritmo

En cada uno de los experimentos descritos en la sección anterior, se utilizó el enfoque de transferencia de aprendizaje. Para ello, cada uno de los cuatro modelos fueron cargados con los pesos de haber sido entrenado con el conocido conjunto de imágenes ImageNet; estos pesos se encuentran disponibles en tensorflow 2.0. Una vez cargados los pesos, cada modelo fue ajustado al conjunto de datos Chest X-Ray Images (*Pneumonia*) de Guangzhou, China.

La tabla-4 describe la cantidad de parámetros totales y entrenables utilizados en cada uno de los modelos propuestos, seguidamente son registrados los valores que definen el consumo computacional del algoritmo como Ram Consumida, Uso en disco y GPU; en cada columna se resalta en rojo el menor valor.

Adicionalmente, como validación del algoritmo y cumplimiento del cuarto objetivo específico, se analizó el comportamiento de cada una de las arquitecturas convolucionales descritas en la fase 3 (3.3) de la metodología, evaluadas a través de la exactitud (*Accuracy*), exhaustividad (*Recall*), precisión (*Pre*), y tiempo en segundos. Estos resultados son mostrados en la tabla-5, en la cual se resalta en rojo el mejor resultado de cada métrica.

Tabla 4. Total de parámetros, parámetros entrenados, Consumo de Ram, disco y GPU, de cada una de las cuatro arquitecturas convolucionales usadas.

Modelo CNN	Parámetros Totales	Parámetros Entrenables	Ram Consumida (GB)	Uso en disco (GB)	GPU Usada (GB)
VGG-16	22,241,990	14,606,726	2.65	32.98	6.60
ResNet-50	48,653,552	48,608,112	2.73	32.98	8.60
Inception-V3	23,851,784	23,817,352	2.75	32.54	8.60
DenseNet-169	14,558,632	14,400,232	2.93	32.68	8.60

Tabla 5. Resultados cuantitativos para cada una de las arquitecturas convolucionales evaluadas.

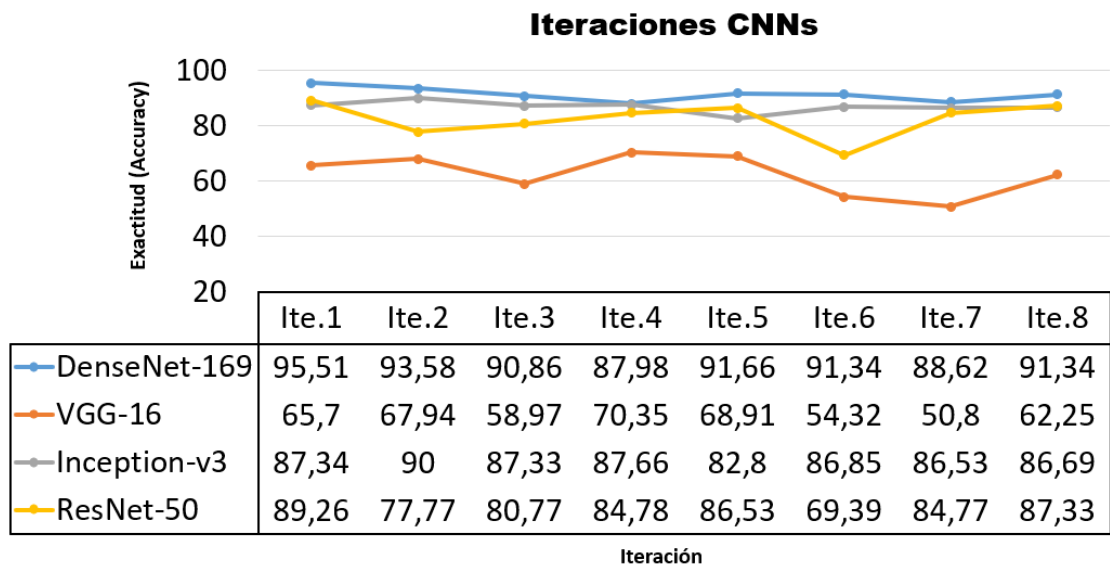
Modelo CNN	<i>Accuracy</i>	<i>Recall</i>	<i>Pre</i>	Tiempo (s)
VGG-16	62.42	50.61	72.21	252.8
ResNet-50	82.5	81.5	82.5	478.2
Inception-V3	86.9	82.5	86.9	239.3
DenseNet-169	91.36	91.12	91.36	356.7

En la tabla-4 y tabla-5 se observa que la arquitectura DenseNet-169 requiere el entrenamiento de un número de parámetros menor que las demás arquitecturas, y presenta menos consumo computacional, al tiempo que permite lograr una tasa de *Accuracy*, *Recall* y *Pre* más alta que con las demás redes evaluadas. Por estos motivos, la red neuronal **DenseNet-169** es seleccionada como modelo convolucional para la clasificación de neumonía a través de imágenes de rayos-X de tórax en este

proyecto, siguiendo la configuración de hiperparámetros presentada en la sección 4.4.

Por otra parte, la figura-22 tabula y grafica el rendimiento de las 8 ejecuciones o iteraciones evaluadas en términos de exactitud (*Accuracy*), para cada una de las redes propuestas, exaltando que el mayor rendimiento lo alcanzó la arquitectura **DenseNet-169** con una exactitud de 95.51 %.

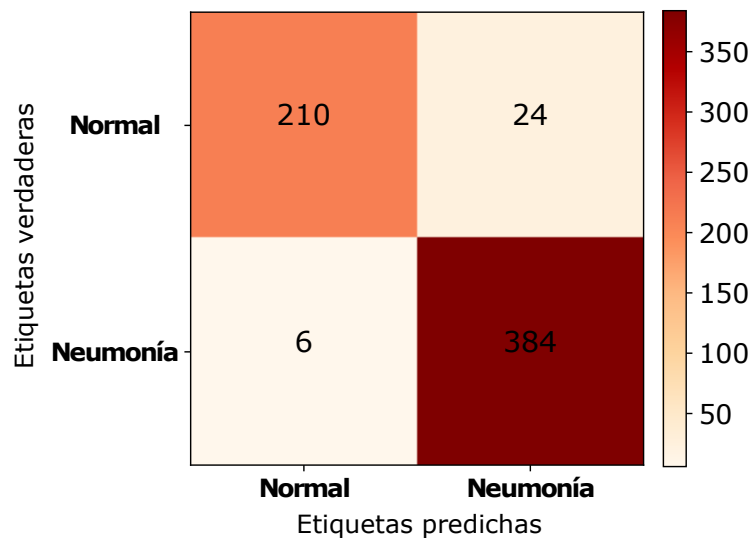
Figura 22. La gráfica tabula el comportamiento de la exactitud respecto a las iteraciones realizadas, ocho en total.



Seguidamente, la figura-23 presenta la matriz de confusión obtenida tras la clasificación de los datos de prueba con la arquitectura elegida (DenseNet-169). La matriz muestra el número de reconocimiento por clase, es decir, de 234 imágenes etiquetadas como normal, 210 fueron etiquetadas correctamente por el algoritmo, mientras que sólo 24 fueron predichas erróneamente; por otra parte, de las 390 imágenes de *test* que corresponde a neumonía, 384 fueron etiquetadas correctamente, y sólo las 6 restantes fueron predichas erróneamente. En general, las imágenes acertadas de la diagonal principal equivalen a las métricas presentadas en la tabla-5, fila

DenseNet-169.

Figura 23. Matriz de confusión de la red DenseNet-169 usando 624 muestras de evaluación.



Finalmente, la figura-24 expone gráficamente el entrenamiento de las cuatro redes neuronales evaluadas: VGG-16, ResNet-50, Inception-V3 y DenseNet-169. La columna de la izquierda presenta la pérdida (*loss*) de cada algoritmo respecto al número de épocas (1 a 10). La tendencia de estas cuatro gráficas ideal es la disminución del valor de pérdida respecto al aumento de épocas, lo cual representa un comportamiento adecuado en la convergencia del modelo. Específicamente, el modelo DenseNet-169, en la fila 4, presenta el mejor desempeño de pérdida respecto a los demás, ya que finaliza con un valor cercano a 0.25, mientras que las demás redes finalizan en valores de 0.4 a 0.5. Nótese también que, a medida que aumenta el número de épocas la pérdida tiende a seguir disminuyendo. Razón por la cual, si se prefiriese un modelo cuyo costo computacional no sea limitado, podría explorarse esta red con mayor número de épocas de entrenamiento.

En general, las curvas de pérdida de los modelos no muestran sobre-entrenamiento, ya que la validación y el entrenamiento en la pérdida se combinan con poca diver-

gencia, decreciendo juntas constantemente. Por otra parte, la columna de la derecha (figura-24), muestra las gráficas de exactitud de la clasificación de la red neuronal en los datos de entrenamiento, a medida que transcurren las épocas (1 a 10). Note que, estas curvas presentan un comportamiento logarítmico, razón por la cual, se definió usar hasta 10 épocas. Este comportamiento es especialmente notable en la arquitectura DenseNet-169, la cual obtiene el comportamiento más estable tanto en validación como prueba, desde el comienzo.

Objetivo específico No. 5. “Comparar los resultados obtenidos con otros métodos similares del estado del arte.”

Como resultado de la quinta fase del trabajo, se expone la recopilación de trabajos relacionados que implementan clasificación binaria para el conjunto de datos presentados en la sección 4.1. **Chest X-Ray Images (Pneumonia)**.

4.6. Revisión del estado del arte.

La tabla-6 muestra la exactitud del Método propuesto en este trabajo, en comparación con otros enfoques en la literatura, en condiciones muy similares.

Tabla 6. Comparación con trabajos similares

Modelo CNN	Accuracy	Recall	Tiempo (s)
TL+DataAug [20]	98	97	-
ResNeTL+DataAug+5CNNs [19]	96.4	99.62	-
TL+DataAug+9CNNs+Norm [15]	96.61	94.92	-
Método Propuesto	91.36	91.12	356.705

Se describe cada uno de los métodos y la cita bibliográfica en la cual se detalla su estrategia de entrenamiento y clasificación, los campos marcados con un guión (–) corresponden a que esta métrica no fue calculada en el documento original. Sin embargo, la autora de este trabajo replicó la implementación de los trabajos

citados en la tabla-6 con el fin de conocer otros datos no reportados en los artículos. Los resultados de la réplica se exponen en la tabla-6, donde se muestra el número de parámetros entrenables de cada método con el cual se comparó este trabajo; además, se expone el número de épocas, el tiempo que tardó cada entrenamiento, y la exactitud alcanzada por el algoritmo replicado en las simulaciones de la autora. Cabe aclarar que, para el primer modelo de la tabla-6 se muestran resultados con 4 números de épocas distintas, debido a que el artículo original no reporta cuántas épocas usaron para entrenar el modelo. Por el contrario, el segundo y tercer método se replicaron con el mismo número de épocas descrito en sus artículos, es decir, 200 y 300 épocas, respectivamente.

Tabla 7. Réplica de trabajos similares

Modelo CNN	Parámetros Total	Épocas	Tiempo(s)	Accuracy
TL+DataAug [20]	20,493,736	10	495.53	67.2
		50	2349.8	78.3
		100	4684.8	86.1
		200	9378.8	89.7
TL+DataAug+5CNNs [19]	94,292,632	200	15270.46	84.9
TL+DataAug+9CNNs+Norm [15]	48,623,344	300	8760.12	85.89

El método TL+DataAug ²⁰ titulado “A novel transfer learning based approach for pneumonia detection in chest X-ray images”(2020), propone un enfoque por aprendizaje de transferencia y aumento de los datos, con el cual alcanzan el 98% de exactitud usando la arquitectura DenseNet-201; el valor más alto según indica la literatura. Sin embargo, este artículo no reporta número de épocas, ni tiempo de entrenamiento. La implementación de el artículo por parte de la autora (ver tabla-7) registró más de 20 millones de parámetros, un tiempo de 9378 segundos para entrenar durante 200 épocas, y una exactitud de 84.9%.

Por otra parte, los autores del trabajo TL+DataAug+5CNNs ¹⁹ titulado “Automated

methods for detection and classification pneumonia based on x-ray images using deep learning"(2020), proponen fusionar modelos de CNNs para clasificar. Ellos usan las arquitecturas: AlexNet, DenseNet121, InceptionV3, ResNet18 y GoogLeNet, para generar un único vector de características, alcanzando una exactitud de un 96.4 %. Sin embargo, estos resultados requieren de entrenar cada arquitectura por 50, 100 y hasta 200 épocas, generando altos costos computacionales. La implementación del artículo realizada por la autora, cuyos resultados se observan en la tabla-7, evidencian un tiempo de entrenamiento de 15270 segundos para 200 épocas, y el valor más alto de parámetros entre estos modelos con 94,292,632 parámetros.

En el trabajo TL+DataAug+9CNNs+Norm ¹⁵ titulado "Automated methods for detection and classification pneumonia based on x-ray images using deep learning"(2020), usan transferencia de aprendizaje, aumento de datos, y técnicas de procesamiento como ajuste fino, normalización y ecualización de histograma, comparan el rendimiento de nueve arquitecturas, y alcanzan una exactitud de 96.61 %, luego de entrenar durante 300 épocas. La implementación de este trabajo por la autora (tabla-7) expone que se alcanzó una exactitud de 85.89 %, utilizando la red con mayor exactitud en el trabajo original: ResNet-50.

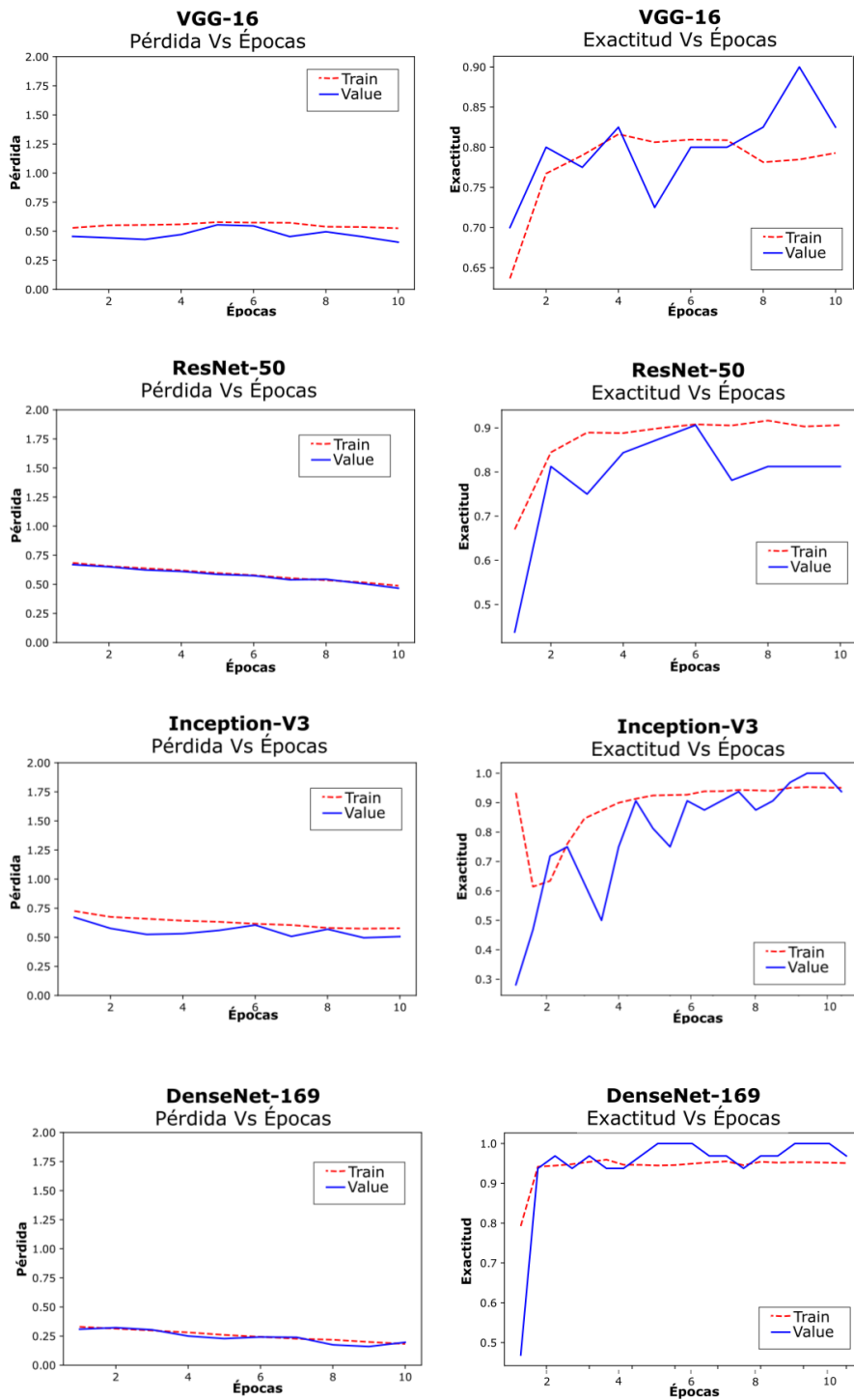
Es de resaltar que, si bien el trabajo ²⁰ obtiene la exactitud más alta del estado del arte, hasta donde sabemos, con 98 %, también ese método alcanzó la exactitud más alta de los trabajos replicados (tabla-7); sin embargo, en la réplica la exactitud fue de 89.7 %. Los demás trabajos, aunque conservaron la relación de exactitud al compararse con los otros métodos, también disminuyeron considerablemente la exactitud en un promedio de 10 % al ser replicados.

La disminución en el rendimiento obtenido al implementar los trabajos con los cuales se comparó el método propuesto, pueden ser atribuidos a varios factores. Por ejemplo, los recursos computacionales con los cuales se imitaron los modelos son considerablemente menores que los usados en los artículos originales. Además, algunos hiperparámetros y detalles más específicos no fueron reportados en los artículos, por lo que leves cambios pudieron influir en la diferencia de exactitud. Incluso en el trabajo ²⁰ que no reporta el número de épocas utilizadas, pudo haber sido usado originalmente un número considerablemente mayor; sin embargo, en aras de guardar proporciones para la comparación se implementaron simulaciones de ese método hasta con 200 épocas.

Finalmente, cabe observar que considerando las tablas-5,6 y 7 el método propuesto requiere de menor tiempo de entrenamiento (356 segundos), menor cantidad de parámetros entrenables (14.4 millones), bajo número de épocas (10), y no implica la concatenación de diferentes arquitecturas para alcanzar resultados de exactitud superiores al 90%. Por lo tanto, podría considerarse como una buena alternativa para ser utilizada por personal que no cuente con equipos de cómputo y tiempo suficientes, como para entrenar modelos de mayor complejidad y costo computacional.

Para recordar, el método propuesto y la réplica de los tres métodos con los cuales se comparó este trabajo fueron ejecutados en equipos con las características computacionales descritas en la sección 2.3.

Figura 24. Curvas de entrenamiento y exactitud, durante el entrenamiento, de las cuatro arquitecturas neuronales convolucionales VGG-16, ResNet-50, Inception-V3 y DenseNet-169.



5. CONCLUSIONES

Este trabajo presentó un enfoque para procesar y clasificar imágenes de rayos-X frontales de tórax en una de dos categorías médicas: normal o con neumonía.

Para este propósito se seleccionó el conjunto de datos Chest X-Ray Images (*Pneumonia*), el cual contiene 5.856 imágenes de rayos-X, adquiridas en el Centro Médico de Niños y Mujeres de Guangzhou, en China. Las imágenes empleadas están compuestas por tres bandas espectrales, y todas han sido etiquetadas por personal médico experto de China, lo cual hace posible entrenar el modelo neuronal de clasificación.

En este trabajo se diseñó una estrategia de tres etapas para etiquetar las imágenes, la cual consiste en Redimensionamiento, Ecuilización y Clasificación con la red neuronal convolucional DenseNet-169.

Además, se presentaron, implementaron y evaluaron cuatro diferentes arquitecturas convolucionales del estado del arte, previamente entrenadas con el conjunto de datos de ImageNet y haciendo uso de la técnica de transferencia de aprendizaje.

Las arquitecturas implementadas fueron evaluadas en términos de tiempo y exactitud. Este trabajo alcanzó una exactitud máxima del 95.51 % y en promedio de 91.36 %, mediante la red neuronal convolucional DenseNet-169. Según las simulaciones realizadas por la autora, el trabajo propuesto permitió una exactitud 1.66 % mayor que los otros métodos, si se comparan bajo las mismas condiciones computacionales. Por otra parte, la exactitud obtenida por el método propuesto usando la red DenseNet-169 es comparable con la obtenida mediante las redes Inception-v3

y ResNet-50; sin embargo, a pesar de su comportamiento similar, estas redes implican mayores costos computacionales debido a la mayor cantidad de parámetros que utilizan. Por lo tanto, según la evaluación, se seleccionó la red DenseNet-169, la cual demostró buen rendimiento en la clasificación en todas las medidas de desempeño.

Por otra parte, este proyecto permitió a la autora aumentar sus destrezas y habilidades de programación, redacción, comprensión del área de inteligencia artificial, y formación integral como estudiante del programa Ingeniería de Sistemas e Informática. El producto principal del proyecto consta de un algoritmo que podría apoyar a los radiólogos en el diagnóstico de neumonía en niños y adultos, inmediatamente después de la adquisición de imágenes frontales de rayos-X.

Se espera que a futuro el desarrollo y perfeccionamiento de algoritmos para la detección de neumonía contribuya a disminuir las brechas de calidad y atención oportuna de los sistemas de atención médica, en especial en las zonas rurales del territorio colombiano.

Trabajo futuro: Finalmente, el desarrollo de este proyecto permitió vislumbrar la importancia de ciertos trabajos futuros a desarrollar: (1) Continuar con la exploración y propuesta de estrategias de clasificación que permitan un alto rendimiento incluso a partir un número reducido de muestras de entrenamiento. (2) Ampliar la capacidad del clasificador para que sea posible etiquetar múltiples infecciones respiratorias (por ejemplo, neumonía bacteriana, neumonía viral, y covid-19), así como incrementar la exactitud de clasificación, manteniendo costos computacionales reducidos.

BIBLIOGRAFÍA

ABADI, Martín, *et al.* *TensorFlow: Large-Scale Machine Learning on Heterogeneous Systems*. Software available from tensorflow.org. 2015 (vid. pág. 47).

— *TensorFlow: Large-Scale Machine Learning on Heterogeneous Systems*. Software available from tensorflow.org. 2015 (vid. pág. 56).

AGUIRRE DOBERNACK, Nicolás. “Implementación De Un Sistema De Detección De Señales De Tráfico Mediante Visión Artificial Basado En FPGA”. En: *Universidad de Sevilla* (2013) (vid. págs. 16, 24).

AKHTAR, Naveed y MIAN, Ajmal. “Threat of adversarial attacks on deep learning in computer vision: A survey”. En: *IEEE Access* 6 (2018), págs. 14410-14430 (vid. pág. 29).

ALPAYDIN, Ethem. *Introduction to machine learning*. MIT press, 2020 (vid. pág. 28).

ARELI, Jamie y BARRERA, Toral. “Redes Neuronales”. [Internet; descargado 24-noviembre-2020] (vid. pág. 31).

ASNAOUI, Khalid El; CHAWKI, Youness e IDRI, Ali. “Automated methods for detection and classification pneumonia based on x-ray images using deep learning”. En: *arXiv preprint arXiv:2003.14363* (2020) (vid. págs. 16, 42, 70).

BAZRAFKAN, Shabab, *et al.* “Deep learning for facial expression recognition: A step closer to a smartphone that knows your moods”. En: *2017 IEEE International Conference on Consumer Electronics (ICCE)*. IEEE. 2017, págs. 217-220 (vid. pág. 39).

BISONG, Ekaba. "Google Colaboratory". En: *Building Machine Learning and Deep Learning Models on Google Cloud Platform*. Springer, 2019, págs. 59-64 (vid. pág. 45).

BUSHBERG, Jerrold T y BOONE, John M. *The essential physics of medical imaging*. Lippincott Williams & Wilkins, 2011 (vid. pág. 25).

CARVAJAL PATIÃ, Daniel Alcides. "Deteccion Automatica Del Nivel De Estratificacion Socioeconomico Urbano Usando Redes Neuronales Convolucionales Sobre Imagenes Satelitales Con Informacion Aumentada." Tesis doct. Universidad Industrial de Santander, Escuela De Ing. De Sistemas, 2018 (vid. pág. 37).

CHOUHAN, Vikash, *et al.* "A novel transfer learning based approach for pneumonia detection in chest X-ray images". En: *Applied Sciences* 10.2 (2020), pág. 559 (vid. págs. 17, 69-71).

CHUNG, Albert Chi Shing, *et al.* "Information processing in medical imaging". En: *Information Processing in Medical Imaging*. 2019 (vid. págs. 21, 22).

DATOS.GOV.CO. *Alcaldía de Bucaramanga - Observatorio Digital Municipal de Bucaramanga-Portal de Datos Abiertos*. 2019 (vid. pág. 13).

DENG, Jia, *et al.* "Imagenet: A large-scale hierarchical image database". En: *2009 IEEE conference on computer vision and pattern recognition*. Ieee. 2009, págs. 248-255 (vid. pág. 51).

DÍAZ, Emili; MARTÍN-LOECHES, Ignacio y VALLÉS, Jordi. "Neumonía nosocomial". En: *Enfermedades Infecciosas y Microbiología Clínica* 31.10 (2013), págs. 692-698 (vid. pág. 19).

DREAMSTIME.COM. *Pulmonía Diferencia y comparación de los bronquiolos de los pulmones y alvéolos y pulmonía sanos*. 2018 (vid. pág. 20).

ESCOBAR, J. P., *et al.* "Transfer Learning with Convolutional Neural Network for Gastrointestinal Diseases Detection using Endoscopic Images". En: *2020 IEEE Colombian Conference on Applications of Computational Intelligence (IEEE ColCACI 2020)*. 2020, págs. 1-6. DOI: 10.1109/ColCACI50549.2020.9247847 (vid. pág. 40).

ESTEVAN, Miguel. "Examen radiográfico del tórax en las neumonías de probable causa bacteriana". En: *Archivos de Pediatría del Uruguay* 73.1 (2002), págs. 15-21 (vid. pág. 14).

FERNÁNDEZ, Luis A; DIAZ, Daniel y DEPAOLI, Roberto. "Optimización de la ecualización del histograma en el procesamiento de imágenes digitales". En: *VII Workshop de Investigadores en Ciencias de la Computación*. 2005 (vid. págs. 16, 26).

FISZMAN, Marcelo, *et al.* "Automatic detection of acute bacterial pneumonia from chest X-ray reports". En: *Journal of the American Medical Informatics Association* 7.6 (2000), págs. 593-604 (vid. pág. 42).

FUENTE CASTILLO, Víctor de la y LOPE ASIAÍN, Javier de. "Diseño evolutivo de arquitecturas de Deep Learning para detección de vías de transporte." En: () (vid. págs. 33, 52).

GAMAGE, C., *et al.* "GI-Net: Anomalies Classification in Gastrointestinal Tract through Endoscopic Imagery with Deep Learning". En: *2019 Moratuwa Engineering Research Conference (MERCOn)*. 2019, págs. 66-71 (vid. pág. 15).

GATTINONI, Luciano, *et al.* *COVID-19 pneumonia: different respiratory treatments for different phenotypes?* 2020 (vid. pág. 16).

GAVILÁN, Ignacio G.R. *Catálogo de componentes de redes neuronales (II): funciones de activación*. [Internet; descargado 19-noviembre-2020]. 2020 (vid. pág. 34).

GELVES MENDOZA, Sergio Alexander y RINCON CADENA, Wanda Catalina. “Metodos De Machine Learning Para El Apoyo Al Diagnostico De Supervivencia De Pacientes Con Sepsis Basadas En Datos Clinicos Y Zimografias.” Tesis doct. Universidad Industrial de Santander, Escuela De Ing. De Sistemas, 2015 (vid. pág. 25).

GOMEZ, Maria y JAIMES, Estefania. “OPTIMIZACION PARAMETROS REDES NEURONALES ARTIFICIALES PREDICCION RENDIMIENTO AGRICOLA CULTIVOS CACAO MEDIANTE METODOS SELECCION INGENIERIA CARACTERISTICAS”. Tesis doct. Universidad Industrial de Santander, Escuela De Ing. De Sistemas, 2019 (vid. pág. 43).

GOMEZ, Natalia Andrea; SANCHEZ, Karen y ARGUELLO, Henry. “Non-Destructive Method for Classification of Cocoa Beans from Spectral Information”. En: *2019 XXII Symposium on Image, Signal Processing and Artificial Vision (STSIVA)*. IEEE. 2019, págs. 1-5 (vid. pág. 56).

GOWER, John C. “A general coefficient of similarity and some of its properties”. En: *Biometrics* (1971), págs. 857-871 (vid. pág. 57).

HE, Kaiming , *et al.* “Deep residual learning for image recognition”. En: *Proceedings of the IEEE conference on computer vision and pattern recognition*. 2016, págs. 770-778 (vid. pág. 50).

HRESCAK, María Cinthya Ortega y SOCOLSKY, Gustavo A. “Godfrey Newbold Hounsfield: historia e impacto de la tomografía computada”. En: *Revista argentina de radiología* 76.4 (2012), págs. 331-341 (vid. págs. 14, 23).

HUANG, Gao, *et al.* “Densely connected convolutional networks”. En: *Proceedings of the IEEE conference on computer vision and pattern recognition*. 2017, págs. 4700-4708 (vid. pág. 52).

HUANG, Jin y LING, Charles X. “Using AUC and accuracy in evaluating learning algorithms”. En: *IEEE Transactions on knowledge and Data Engineering* 17.3 (2005), págs. 299-310 (vid. pág. 56).

INTERNET. *Introducción a las imágenes digitales*. 2018 (vid. pág. 26).

— *Redes neuronales en inteligencia artificial*. 2019 (vid. pág. 30).

INTERNET, OspreyData. *AI/ML MODELS 101: WHAT IS A MODEL?* 1997 (vid. pág. 28).

IZAURIETA, Fernando y SAAVEDRA, Carlos. “Redes neuronales artificiales”. En: *Departamento de Física, Universidad de Concepción Chile* (2000) (vid. pág. 32).

KADOYA, Noriyuki, *et al.* “Evaluation of four-dimensional cone beam computed tomography ventilation images acquired with two different linear accelerators at various gantry speeds using a deformable lung phantom”. En: *Physica Medica* 77 (2020), págs. 75-83 (vid. pág. 24).

KARTHIKZIFFER. *Keras Densenet CIFAR10*. 2019 (vid. pág. 53).

KERMANY, Daniel S, *et al.* “Identifying medical diagnoses and treatable diseases by image-based deep learning”. En: *Cell* 172.5 (2018), págs. 1122-1131 (vid. pág. 15).

KETCHAM, Richard A y CARLSON, William D. "Acquisition, optimization and interpretation of X-ray computed tomographic imagery: applications to the geosciences". En: *Computers & Geosciences* 27.4 (2001), págs. 381-400 (vid. pág. 24).

KHAN, SanaUllah, *et al.* "A novel deep learning based framework for the detection and classification of breast cancer using transfer learning". En: *Pattern Recognition Letters* 125 (2019), págs. 1-6 (vid. pág. 15).

KHANDELWAL, Vaibhav. *The Architecture and Implementation of VGG-16*. 2020 (vid. pág. 49).

LARRANAGA, Pedro; INZA, Inaki y MOUJAHID, Abdelmalik. "Tema 8. redes neuronales". En: *Redes Neuronales, U. del P. Vasco* 12 (1997), pág. 17 (vid. pág. 34).

LECUN, Yann; BENGIO, Yoshua e HINTON, Geoffrey. "Deep learning". En: *nature* 521.7553 (2015), págs. 436-444 (vid. pág. 30).

LEE, Honglak, *et al.* "Unsupervised feature learning for audio classification using convolutional deep belief networks". En: *Advances in neural information processing systems*. 2009, págs. 1096-1104 (vid. pág. 42).

LIANG, Xuezhi, *et al.* "Soft-margin softmax for deep classification". En: *International Conference on Neural Information Processing*. Springer. 2017, págs. 413-421 (vid. pág. 30).

LITJENS, Geert, *et al.* "A survey on deep learning in medical image analysis". En: *Medical image analysis* 42 (2017), págs. 60-88 (vid. págs. 14, 42).

LÓPEZ, Raquel Florez; FERNANDEZ, Jose Miguel y FERNÁNDEZ, José Miguel Fernández. *Las redes neuronales artificiales*. Netbiblo, 2008 (vid. pág. 32).

LOVERA, Dolores, *et al.* “Neumonía en lactantes de 6 meses de edad y menores”. En: *Pediatría (Asunción)* 47.2 (2020), págs. 74-80 (vid. pág. 19).

MALACARA, Daniel. *Óptica tradicional y moderna*. Sirsi) i9789681666187. 2002 (vid. pág. 26).

MARGARET, Rouse. *Aprendizaje profundo (deep learning)*. 2020 (vid. pág. 29).

MARTINEZ, Fabio. *ai-2-UIS-student*. 2020 (vid. pág. 36).

MARTINEZ, Hani y EDUARDO, Rafael. “Deteccion De Danos En Puentes De Tipologia Viga Y Losa Simplemente Apoyada Usando Redes Neuronales Artificiales.” Tesis doct. Universidad Industrial de Santander, Escuela De Ing. Civil, 2012 (vid. pág. 43).

MARTÍNEZ, J. *Factores asociados a neumonía adquirida en comunidad en infantes menores de 5 años, Subcentro de Salud no. 1, Cuenca*. 2015 (vid. pág. 41).

MDSPD, Ecuador. “Neumonía adquirida en la comunidad en pacientes de 3 meses a 15 años”. En: *Guía de Práctica Clínica* 13 (2017) (vid. pág. 41).

MINSALUD. *Infecciones Respiratorias Agudas - Ministerio de Salud y protección Social*. 2020 (vid. pág. 13).

MISSINGLINK.AI. *Fully Connected Layers in Convolutional Neural Networks: The Complete Guide*. 2018 (vid. pág. 37).

— *The Complete Guide to Artificial Neural Networks: Concepts and Models*. 2018 (vid. págs. 31, 35).

MITCHELL, Tom M, *et al.* *Machine learning*. 1997 (vid. pág. 28).

MONDACA, Roberto. "Por qué reducir las dosis de radiación en pediatría". En: *Revista chilena de radiología* 12.1 (2006), págs. 28-32 (vid. pág. 23).

MOONEY, Paul. *Chest X-Ray Images (Pneumonia)*. 2018 (vid. págs. 59, 60, 62).

NA8. *Aprendizaje Profundo: una Guía rápida*. [Internet; descargado 24-noviembre-2020]. 2017 (vid. pág. 32).

NIELSEN, Michael A. *Neural networks and deep learning*. Vol. 2018. Determination press San Francisco, CA, 2015 (vid. págs. 34, 37).

NORTH AMERICA (RSNA), Radiological Society of. *Seguridad del paciente: Dosis de radiación en exámenes de rayos X y TAC Qué son los rayos X y qué hacen*. 2013 (vid. pág. 22).

OMS. *Organización Mundial de Salud-Neumonía*. 2019 (vid. pág. 13).

ORTIZ SUAREZ, Juan Manuel. "Evaluacion De Metodos Con Redes Neuronales Convolucionales Para La Clasificacion De La Enfermedad De Alzheimer En Imagenes De Resonancia Magnetica Cerebral." Tesis doct. Universidad Industrial de Santander, Escuela De Ing. De Sistemas, 2016 (vid. pág. 43).

PAN, Sinno Jialin y YANG, Qiang. "A survey on transfer learning". En: *IEEE Transactions on knowledge and data engineering* 22.10 (2009), págs. 1345-1359 (vid. pág. 17).

PÉREZ, P. y VALENTE, M. *Fundamentos básicos del procesamiento de imágenes*. 2018 (vid. págs. 16, 26).

RAHMAN, Tawsifur, *et al.* “Transfer Learning with Deep Convolutional Neural Network (CNN) for Pneumonia Detection using Chest X-ray”. En: *Applied Sciences* 10.9 (2020), pág. 3233 (vid. págs. 17, 42, 69).

RAJPURKAR, Pranav, *et al.* “Chexnet: Radiologist-level pneumonia detection on chest x-rays with deep learning”. En: *arXiv preprint arXiv:1711.05225* (2017) (vid. págs. 15, 42).

RAMIREZ RUIZ, Jose Mauricio. “Caracterizacion De Niños Entre 2 Meses Y 17 Años Con Diagnostico De Neumonia Adquirida En La Comunidad Hospitalizados En El Hus.” Tesis doct. Universidad Industrial de Santander, Escuela de Medicina, Departamento De . . . , 2016 (vid. pág. 41).

SAHINER, Berkman, *et al.* “Deep learning in medical imaging and radiation therapy”. En: *Medical physics* 46.1 (2019), e1-e36 (vid. pág. 21).

SAMIA. *Towards Machine Learning* (vid. pág. 37).

SELF, Wesley H, *et al.* “High discordance of chest x-ray and computed tomography for detection of pulmonary opacities in ED patients: implications for diagnosing pneumonia”. En: *The American journal of emergency medicine* 31.2 (2013), págs. 401-405 (vid. pág. 41).

SHAH, Kanish. *Una descripción general rápida de la transferencia de aprendizaje y su importancia en las aplicaciones del mundo real.* 2020 (vid. pág. 40).

SIMONYAN, Karen y ZISSERMAN, Andrew. “Very deep convolutional networks for large-scale image recognition”. En: *arXiv preprint arXiv:1409.1556* (2014) (vid. pág. 50).

SONG, Fengxiang, *et al.* “Emerging 2019 novel coronavirus (2019-nCoV) pneumonia”. En: *Radiology* 295.1 (2020), págs. 210-217 (vid. pág. 16).

TAKEMBO, Clovis Ntahkie, *et al.* “Effect of electromagnetic radiation on the dynamics of spatiotemporal patterns in memristor-based neuronal network”. En: *Nonlinear Dynamics* 95.2 (2019), págs. 1067-1078 (vid. pág. 14).

URIBARRI, Salvador Galindo. “Principios matemáticos de la reconstrucción de imágenes tomográficas”. En: *CIENCIA ergo-sum* 10.3 (2003), págs. 271-281 (vid. pág. 23).

VAN DYK, David A y MENG, Xiao-Li. “The art of data augmentation”. En: *Journal of Computational and Graphical Statistics* 10.1 (2001), págs. 1-50 (vid. pág. 47).

WANG, Cheng, *et al.* “Pulmonary image classification based on inception-v3 transfer learning model”. En: *IEEE Access* 7 (2019), págs. 146533-146541 (vid. pág. 51).

WIKIPEDIA. *Propagación hacia atrás — Wikipedia, La enciclopedia libre*. [Internet; descargado 19-noviembre-2020]. 2020 (vid. pág. 35).

WIKIPEDIA CONTRIBUTORS. *ImageNet — Wikipedia, The Free Encyclopedia*. [Online; accessed 27-November-2020]. 2020 (vid. pág. 38).

WILLEMINK, Martin J, *et al.* “Preparing medical imaging data for machine learning”. En: *Radiology* 295.1 (2020), págs. 4-15 (vid. pág. 22).

WONG, S. C., *et al.* “Understanding Data Augmentation for Classification: When to Warp?” En: *2016 International Conference on Digital Image Computing: Techniques and Applications (DICTA)*. 2016, págs. 1-6. DOI: 10.1109/DICTA.2016.7797091 (vid. pág. 47).

WONG, Sebastien C , *et al.* "Understanding data augmentation for classification: when to warp?" En: *2016 international conference on digital image computing: techniques and applications (DICTA)*. IEEE. 2016, págs. 1-6 (vid. pág. 47).

WORLDHEALTHORGANIZATION , *et al.* *Standardization of interpretation of chest radiographs for the diagnosis of pneumonia in children*. Inf. téc. World Health Organization, 2001 (vid. pág. 24).

ANEXOS

Anexo A. Productos Académicos

Durante los estudios de Ingeniería de Sistemas e Informática la autora del presente proyecto obtuvo los siguientes productos académicos.

- Gomez, N. A., Sanchez, K., & Arguello, H. (2019, April). Non-Destructive Method for Classification of Cocoa Beans from Spectral Information. In 2019 XXII Symposium on Image, Signal Processing and Artificial Vision (STSIVA) (pp. 1-5). IEEE.

- Escobar, J. P., Gomez, N., Sanchez, K., & Arguello, H. (2020, August). Transfer Learning with Convolutional Neural Network for Gastrointestinal Diseases Detection using Endoscopic Images. In 2020 IEEE Colombian Conference on Applications of Computational Intelligence (IEEE ColCACI 2020) (pp. 1-6). IEEE.