



Proyecto 2:

“Aplicación de la Visión Artificial para la detección de neumonía mediante análisis de imágenes para la salud pública”.

Felipe Pereira Alarcón NRC 7056.

Agustín Muñoz Pino NRC 7055.

Profesor Guía: Dr. Billy Peralta Márquez.

Providencia, Chile

2023.

Tabla de contenido

PRÓLOGO	I
1. INTRODUCCIÓN	4
1.1 Trasfondo	4
1.2 Motivación	4
1.3 Objetivos	5
2. DEFINICIÓN DE PROYECTO	6
2.1 Definición Formal.....	6
2.2 Importancia del Proyecto	6
2.3 Contexto del Proyecto	7
3. DESCRIPCIÓN DE PROYECTO	8
3.1 Descripción del dataset	8
3.2 Resumen	10
4. EXPERIMENTACIÓN DEL PROYECTO	11
4.1 Diseño de Experimentos.....	11
4.2 Enfoque	11
4.3 Métricas de validación propuestas	12
5. RESULTADOS DEL PROYECTO	13
5.1 Reporte de resultados	13

Tabla de Figuras

Figura 1: Tabla Información Dataset.....	8
Figura 2: Información división dataset en conjuntos	9
Figura 3: Imágenes dataset Training.....	9
Figura 4: Información entrenamiento modelo	13
Figura 5: Gráfico de perdida respecto a Épocas	14
Figura 6: Vista previa Imágenes del entrenamiento.....	14
Figura 7: Vista Previa Imágenes posterior al entrenamiento	15

CAPITULO 1

INTRODUCCIÓN

1.1 Trasfondo

En los últimos años, la visión artificial ha surgido como una disciplina de investigación y aplicación tecnológica que tiene como objetivo emular la capacidad visual humana en sistemas computacionales.

La visión artificial se basa en la extracción de información visual de imágenes o secuencias de imágenes, su posterior análisis, preprocesamiento y procesamiento para obtener conocimientos significativos y facilitar la toma de decisiones inteligentes.

Esta disciplina ha encontrado numerosas aplicaciones en diversos campos, como la industria, la robótica, la medicina y la seguridad. En la industria, por ejemplo, se utiliza para inspeccionar productos, controlar la calidad y automatizar procesos.

En el ámbito médico, se emplea para el diagnóstico y la detección temprana de enfermedades basándose en imágenes médicas. En términos de seguridad, se utiliza para la vigilancia y la identificación de objetos o comportamientos sospechosos.

El progreso de la visión artificial ha sido impulsado por diferentes factores. En primer lugar, los avances en hardware, como el incremento en la capacidad de procesamiento y el desarrollo de sensores de alta resolución, han permitido adquirir y analizar imágenes con mayor precisión y rapidez.

Además, los algoritmos y técnicas de aprendizaje automático, como las redes neuronales convolucionales, han demostrado un gran potencial para el reconocimiento de objetos y patrones en imágenes.

1.2 Motivación

En este informe, nos enfocaremos en la exploración y el desarrollo de soluciones en el campo de la visión artificial. Nuestro objetivo principal consistirá en diseñar y entrenar modelos de visión por computadora capaces de llevar a cabo tareas específicas, como la clasificación de objetos, la detección de anomalías o el seguimiento de objetos en tiempo real.

Para lograrlo, haremos uso de técnicas de aprendizaje automático (Machine Learning) y redes neuronales convolucionales (CNN), aprovechando la capacidad de procesamiento paralelo y la capacidad de generalización que ofrecen estos enfoques.

Además, abordaremos desafíos importantes asociados con la visión artificial, como la variabilidad de las condiciones de iluminación, la presencia de ruido y la necesidad de realizar procesamiento en tiempo real.

Para superar estos desafíos, emplearemos métodos de preprocesamiento de imágenes, técnicas de aumento de datos y optimizaciones de algoritmos.

1.3 Objetivos

El objetivo principal de este proyecto es conseguir una automatización segura a través de técnicas de machine learning aplicadas en este proyecto, para asegurar un mejor análisis y dar un diagnóstico cercano a lo definitivo de neumonía a través de visión artificial.

Continuando, se plantea implementar un modelo de visión artificial (CV) utilizando una red neuronal convolucional (CNN) basada en la arquitectura ResNet-50, pre-entrenada en el conjunto de datos ImageNet, de manera que el análisis de las imágenes del dataset sea analizada con mejor precisión y asegurar una predicción más acertada.

Aplicar un entrenamiento adecuado para el modelo utilizando el conjunto de datos de entrenamiento del dataset y optimizar los parámetros del modelo utilizando el optimizador Adam (el más utilizado ya que da mejores resultados) y la función de pérdida (Criterio) “Cross-Entropy”.

Evaluar el desempeño del modelo en el conjunto de datos de validación del dataset y calcular su precisión y su pérdida, junto a su tiempo de ejecución.

Los resultados de esta investigación contribuirán al avance de la visión artificial y su aplicación en diversos campos principalmente en la medicina, ofreciendo soluciones eficientes y efectivas para el análisis de imágenes y la toma de decisiones automatizadas.

A través de la exploración de diferentes enfoques y la evaluación de su rendimiento, esperamos mejorar la precisión y la robustez de los sistemas de visión artificial, abriendo así nuevas oportunidades en ámbitos como la medicina en este caso, la seguridad, la industria y muchos otros.

CAPÍTULO 2

DEFINICIÓN DE PROYECTO

2.1 Definición Formal

La problemática planteada consiste en la clasificación de imágenes médicas de rayos X en dos clases: neumonía y normales. el objetivo es desarrollar un modelo de visión artificial capaz de identificar de manera precisa y automatizada la presencia de neumonía en las imágenes.

La detección temprana de neumonía en imágenes de rayos X de tórax es de suma importancia en el ámbito médico, ya que puede facilitar un diagnóstico rápido y preciso, lo que a su vez puede mejorar los resultados del tratamiento y reducir el riesgo de complicaciones para los pacientes.

La dificultad radica en la naturaleza compleja de las imágenes médicas y en la variabilidad de las características presentes en las radiografías de neumonía. Además, las imágenes pueden contener ruido, artefactos y variaciones en las condiciones de iluminación, lo que dificulta la extracción de características relevantes para la clasificación precisa.

Por tanto, la problemática implica el desarrollo de un modelo de visión artificial capaz de aprender patrones y características distintivas de las imágenes de rayos X de tórax para diferenciar entre casos de neumonía y casos normales. Esto implica la necesidad de diseñar una arquitectura de red neuronal convolucional adecuada, en este caso ResNet50, realizar un procesamiento adecuado de las imágenes y utilizar técnicas de aprendizaje automático para entrenar y optimizar el modelo.

El abordaje exitoso de esta problemática tiene el potencial de mejorar la detección y diagnóstico temprano de la neumonía, permitiendo una intervención médica oportuna y mejorando los resultados clínicos. Además, el desarrollo de técnicas de visión artificial para la clasificación de imágenes médicas puede abrir puertas a aplicaciones más amplias en el campo de la medicina y la salud, brindando herramientas precisas y automatizadas para la detección y diagnóstico de diversas enfermedades y condiciones médicas.

2.2 Importancia del Proyecto

El proyecto de clasificación de imágenes médicas de rayos X de tórax utilizando visión artificial tiene una gran importancia en el campo de la salud y la medicina. La detección temprana y precisa de enfermedades es crucial para el tratamiento efectivo y la mejora de los resultados clínicos. En el caso particular de la neumonía, una enfermedad pulmonar común y potencialmente grave, su identificación oportuna puede marcar la diferencia en la vida de los pacientes.



La aplicación de técnicas de visión artificial en el análisis de imágenes médicas ofrece una serie de beneficios significativos. En primer lugar, permite realizar una evaluación rápida y automatizada de un gran volumen de imágenes del dataset, lo que ahorra tiempo y recursos en comparación con los métodos tradicionales de análisis manual. Esto es especialmente importante en entornos clínicos donde el tiempo es crucial y se requiere una toma de decisiones rápida.

Además, la visión artificial puede ayudar a mejorar la precisión y objetividad en el diagnóstico médico. Al entrenar modelos de aprendizaje automático para reconocer patrones y características específicas en las imágenes de rayos X de tórax, se pueden obtener resultados consistentes y confiables. Esto reduce el riesgo de errores humanos y proporciona una base sólida para la toma de decisiones clínicas.

Otro aspecto importante por observar es la posibilidad de implementar sistemas de detección y diagnóstico asistidos por computadora que puedan funcionar como una segunda opinión para los médicos. Estos sistemas pueden ayudar a identificar casos sospechosos, resaltar regiones de interés, evitar falsos positivos y proporcionar información adicional que respalde el proceso de diagnóstico. Esto permite una evaluación más completa y precisa, mejorando la calidad de la atención médica.

Además de la importancia clínica, el proyecto también contribuye al avance de la visión artificial y el aprendizaje automático en general. Al enfrentar el desafío de clasificar imágenes médicas complejas, se pueden desarrollar y perfeccionar técnicas y algoritmos específicos que luego pueden aplicarse a otras áreas y problemas.

2.3 Contexto del Proyecto

La visión artificial ha surgido como una disciplina de investigación y aplicación tecnológica que busca emular la capacidad visual humana en sistemas computacionales. En el campo de la medicina, la visión artificial ha demostrado un gran potencial para mejorar el diagnóstico y la detección temprana de enfermedades a través del análisis de imágenes médicas, como los rayos X de tórax.

El diagnóstico preciso de enfermedades pulmonares, como la neumonía, es fundamental para ofrecer un tratamiento efectivo y mejorar los resultados clínicos. Sin embargo, el análisis manual de un gran número de imágenes radiográficas es un proceso laborioso y propenso a errores humanos. En este contexto, la aplicación de técnicas de visión artificial puede ser de gran ayuda.

La visión artificial ofrece la posibilidad de automatizar el análisis de imágenes radiográficas, permitiendo una evaluación rápida y precisa. Al entrenar modelos de aprendizaje automático con una amplia variedad de imágenes de rayos X de tórax, es posible desarrollar algoritmos capaces de reconocer patrones y características específicas asociadas a la presencia de neumonía.

El proyecto tiene como objetivo principal desarrollar un sistema de clasificación de imágenes médicas de rayos X de tórax utilizando visión artificial. Para ello, se utilizarán técnicas de procesamiento de imágenes y redes neuronales convolucionales, que son modelos de aprendizaje automático especialmente diseñados para el análisis de imágenes.

La clasificación precisa y automatizada de las imágenes de rayos X permitirá identificar de manera eficiente los casos de neumonía, facilitando así la labor de los profesionales de la salud y acelerando el proceso de diagnóstico. Esto puede tener un impacto significativo en la calidad de la atención médica, ya que permitirá tomar decisiones clínicas fundamentadas en una base objetiva y confiable.

Además, este proyecto contribuirá al avance de la visión artificial en el campo de la medicina. Al enfrentar el desafío específico de clasificar imágenes radiográficas de rayos X de tórax, se podrán explorar y perfeccionar técnicas y algoritmos específicos que posteriormente podrán aplicarse a otras áreas médicas y problemas similares.

CAPÍTULO 3

DESCRIPCIÓN DE PROYECTO

3.1 Descripción del dataset

A continuación, se presenta el dataset implementado para este proyecto, el cuál lo puedes descargar desde la página oficial de MedMNIST: <https://medmnist.com/>

Según la fuente "<https://medmnist.com/>", el dataset se conforma de:

Dataset	Modalidad de los datos	Clasificación	Muestras	Entrenamiento (training)	Validación (validation)	Pruebas (test)
PneumoniaMNIST	Tórax X-ray	Clase Binaria	5856	4708	524	624

Figura 1: Tabla Información Dataset

Donde la cantidad de imágenes totales son de 5856 muestras dividiéndose en:

- 4708 imágenes pertenecen al entrenamiento.
- 524 imágenes pertenecen a la validación.
- 624 imágenes pertenecen a las pruebas.

Esto queda demostrado en código ejecutando la casilla 7 del cuaderno de Google Collab:

```
Número de imágenes en el conjunto de entrenamiento: 4708
Número de imágenes en el conjunto de validación: 524
Número de imágenes en el conjunto de prueba: 624
```

Figura 2: Información división dataset en conjuntos

Las imágenes tienen una resolución de 28 píxeles x 28 píxeles en un canal de color basado de 0 a 255, ya que fueron re-escaladas de tamaños más grandes..

Esto queda demostrado en Código ejecutando la casilla 7 del cuaderno de Google Collab:

Description: The PneumoniaMNIST is based on a prior dataset of 5,856 pediatric chest X-Ray images. The task is binary-class classification of pneumonia against normal. We split the source training set with a ratio of 9:1 into training and validation set and use its source validation set as the test set. **The source images are gray-scale, and their sizes are (384-2,916) x (127-2,713). We center-crop the images and resize them into 1x28x28.**

License: CC BY 4.0

Las imágenes están en calidad Grayscale (Escala de grises), lo que supone una dificultad mayor al usar técnicas de visión estándar para imágenes como lo es ResNet50, que están hechas para imágenes RGB (3 canales).

Para revisar las 100 primeras imágenes del dataset, debes ejecutar la casilla 8 del cuaderno de Google Collab:



Figura 3: Imágenes dataset Training

Las cantidades de clases a utilizar son dos según la fuente “<https://medmnist.com/>”, ya que solo pueden existir dos diagnósticos posibles (Binary-Class), los cuales son:

- **Neumonía:** Se clasifican en esta clase, aquellos casos que fueron diagnosticados por el modelo como posible neumonía.
- **Normales:** Se clasifican en esta clase, aquellos casos que fueron diagnosticados por el modelo como sanos.

3.2 Resumen

El proyecto se enfoca en el estudio y aplicación de la visión artificial como una herramienta eficiente para el análisis y detección de neumonía a través de imágenes médicas. El objetivo principal de este proyecto es explorar y evaluar el potencial de la visión artificial en el ámbito de la salud pública, específicamente en la detección temprana de neumonía.

Para lograr dicho objetivo, se utiliza un dataset especialmente creado y adaptado para esta investigación por la biblioteca MedMNIST. Este dataset está compuesto por una amplia variedad de imágenes médicas, obtenidas de diferentes pacientes con características clínicas y patológicas asociadas a la neumonía.

Estas imágenes son procesadas y analizadas mediante técnicas avanzadas de visión artificial, tales como el reconocimiento de patrones, extracción de características y aprendizaje automático (Resnet50).

Los hallazgos obtenidos en este proyecto demuestran que la visión artificial puede desempeñar un papel crucial en la detección temprana de neumonía, permitiendo a los profesionales de la salud tomar decisiones más rápidas y precisas en el diagnóstico de esta enfermedad.

Además, se discuten las implicaciones y posibles aplicaciones de esta tecnología en el ámbito de la salud pública, como la implementación de sistemas de detección automatizados en centros de atención médica y la mejora de la eficiencia en los programas de detección de enfermedades respiratorias.

EXPERIMENTACIÓN DEL PROYECTO

4.1 Diseño de Experimentos

Los experimentos planteados para este proyecto se dividen en:

- Variar los hiperparámetros del modelo, como el número de épocas de 20 a 10, el tamaño del lote del batch_size y la tasa de aprendizaje pasando de 0.1 a 0.001 (Learning rate), y observar cómo afectan al rendimiento del modelo en los conjuntos de validación y prueba.
- Realizar diferentes divisiones de los datos en conjuntos de entrenamiento, validación y prueba y evaluar cómo se ve afectado el rendimiento del modelo solo usando el conjunto de entrenamiento sin las validaciones, y luego implementándole los dos últimos conjuntos indicados anteriormente.
- Probar diferentes arquitecturas de modelos, como utilizar técnicas de visión estándar tales como ResNet18 o auto-keras a ResNet 50, usar una red neuronal convolucional diferente o agregar capas adicionales al modelo existente, y comparar su rendimiento con el modelo base.
- Utilizar técnicas de aumento de datos para aumentar la cantidad de datos de entrenamiento y mejorar el rendimiento del modelo.

4.2 Enfoque

El enfoque por utilizar en este proyecto es el enfoque hold-out, el cual es una técnica de evaluación de modelos que se utiliza para medir su rendimiento y generalización.

En este enfoque, el conjunto de datos disponible se divide en tres conjuntos distintos: un conjunto de entrenamiento, un conjunto de validación y un conjunto de prueba, que luego pasan a ser evaluados en dos módulos, el módulo uno se encarga de entrenar el conjunto de entrenamiento y el módulo dos se encarga de validar y testear los datos usando los conjuntos de validación y test.

El conjunto de entrenamiento se utiliza para entrenar el modelo de IA, es decir, se utiliza para ajustar los pesos y parámetros del modelo mediante algoritmos de aprendizaje automático. Por otro lado, el conjunto de prueba se utiliza para evaluar el rendimiento del modelo después de haber sido entrenado.

La división entre los conjuntos de entrenamiento y prueba debe hacerse de manera aleatoria y asegurarse de que los datos de prueba no se utilicen durante el entrenamiento del modelo. Esto es importante para obtener una evaluación objetiva del rendimiento del modelo en datos no vistos previamente.

Luego de explicar todo esto, en nuestro modelo del proyecto implementamos el 70% del conjunto de entrenamiento, el 15% para la validación y el 15% para testeo.

4.3 Métricas de validación propuestas

Las métricas de validación propuestas en este proyecto son:

- **Precisión (accuracy):** La precisión se calcula dividiendo el número de predicciones correctas por el número total de muestras en el conjunto de datos de validación. Esta métrica muestra la proporción de muestras clasificadas correctamente por el modelo.

El código calcula la precisión en el conjunto de validación utilizando la siguiente línea de código de la casilla 12 del cuaderno de Google Collab:

$$\text{accuracy_val} = \text{total_correct_val} / \text{total_samples_val}$$

donde `accuracy_val` almacena la precisión del modelo en el conjunto de validación.

- **Pérdida (loss):** La pérdida se calcula utilizando la función de pérdida cruzada (`CrossEntropyLoss`) entre las salidas predichas por el modelo y las etiquetas reales en el conjunto de datos de validación. Esta métrica indica cuán cerca están las predicciones del modelo de las etiquetas reales.

El código calcula la pérdida promedio en cada época utilizando la siguiente línea de código de la casilla 12 del cuaderno de Google Collab:

$$\text{epoch_loss} = \text{running_loss} / \text{len}(\text{train_dataset})$$

donde `epoch_loss` almacena la pérdida promedio en el conjunto de entrenamiento durante una época determinada.

5.1 Reporte de resultados

A continuación, se mostrarán los resultados del entrenamiento del modelo luego de analizar el dataset, realizar el modelo con modificaciones de la técnica de visión estándar ResNet-50, para la adaptabilidad de imágenes Grayscale y realizada una simulación de 3 canales RGB en su debido uso para su posterior predicción de Diagnósticos de neumonía.

- **Precisión (Accuracy):** La precisión es una métrica comúnmente utilizada para evaluar modelos de clasificación. Representa la proporción de predicciones correctas en relación con el total de muestras. En tu código, se calcula la precisión en el conjunto de validación y el conjunto de prueba. Un valor alto de precisión indica un buen rendimiento del modelo.
- **Pérdida (Loss):** La pérdida es una métrica que representa qué tan bien se ajustan las predicciones del modelo a las etiquetas reales durante el entrenamiento. En cada época, se registra la pérdida promedio en el conjunto de entrenamiento. Un valor bajo de pérdida indica un mejor ajuste del modelo a los datos de entrenamiento.

En la figura se puede divisar el entrenamiento por épocas (el total de épocas son 10), la perdida se mantiene muy baja y el tiempo es considerablemente rápido para su debido análisis.

```
Época 1/10, Pérdida: 0.2102, Tiempo: 86.28 segundos
Precisión en el conjunto de validación: 40.50763358778626
Época 2/10, Pérdida: 0.1352, Tiempo: 79.96 segundos
Precisión en el conjunto de validación: 37.01908396946565
Época 3/10, Pérdida: 0.1028, Tiempo: 80.82 segundos
Precisión en el conjunto de validación: 38.61832061068702
Época 4/10, Pérdida: 0.0973, Tiempo: 80.39 segundos
Precisión en el conjunto de validación: 39.06488549618321
Época 5/10, Pérdida: 0.0754, Tiempo: 81.18 segundos
Precisión en el conjunto de validación: 38.545801526717554
Época 6/10, Pérdida: 0.0589, Tiempo: 80.91 segundos
Precisión en el conjunto de validación: 39.79007633587786
Época 7/10, Pérdida: 0.0684, Tiempo: 80.78 segundos
Precisión en el conjunto de validación: 39.32442748091603
Época 8/10, Pérdida: 0.0590, Tiempo: 80.95 segundos
Precisión en el conjunto de validación: 39.05725190839695
Época 9/10, Pérdida: 0.0532, Tiempo: 80.90 segundos
Precisión en el conjunto de validación: 39.69083969465649
Época 10/10, Pérdida: 0.0447, Tiempo: 81.11 segundos
Precisión en el conjunto de validación: 38.33587786259542
Precisión en el conjunto de prueba: 34.24679487179487
```

Figura 4: Información entrenamiento modelo

- **Gráfico de pérdida por época:** El gráfico de pérdida por época muestra cómo la pérdida del modelo cambia a lo largo del tiempo. Esta entrega información sobre la convergencia del modelo durante el entrenamiento.

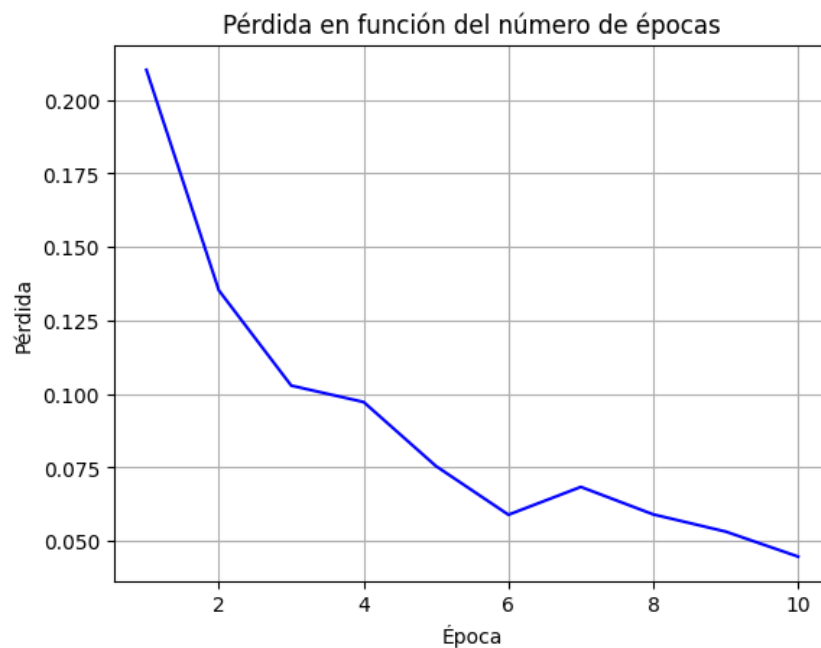


Figura 5: Gráfico de perdida respecto a Épocas

- **Visualización de imágenes:** El código muestra una muestra de 100 imágenes del conjunto de entrenamiento. Se mostrarán imágenes para verificar visualmente si el modelo está clasificando correctamente las imágenes de neumonía.

Imágenes antes del análisis:



Figura 6: Vista previa Imágenes del entrenamiento

Imágenes después del análisis:

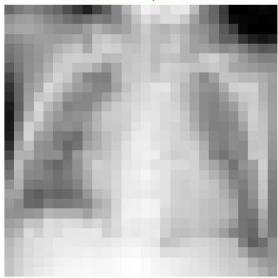
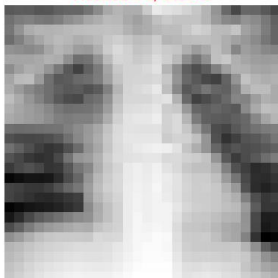
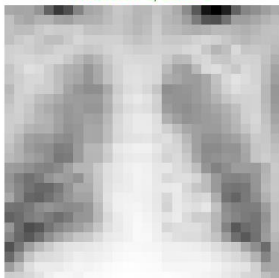
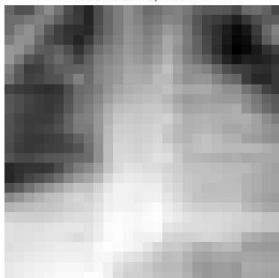

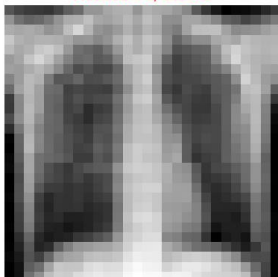
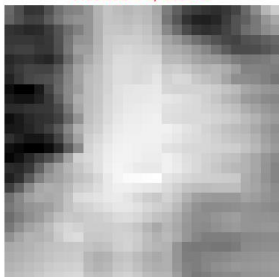
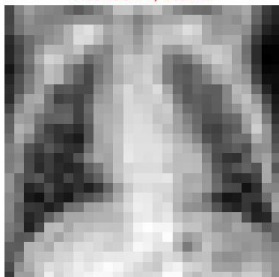
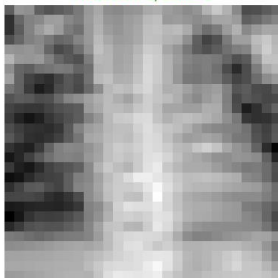
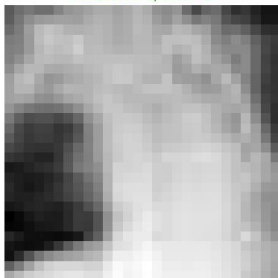
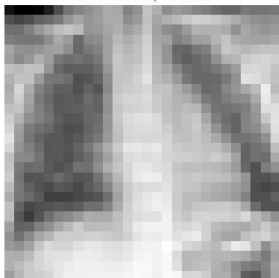
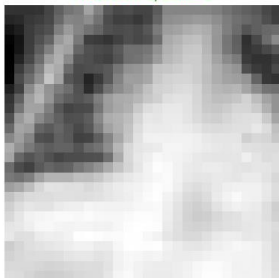
<p>Predicción: 0, Clase: 0</p>  <p>Train_dataset[0]</p>	<p>Predicción: 0, Clase: 1</p>  <p>Train_dataset[1]</p>	<p>Predicción: 1, Clase: 1</p>  <p>Train_dataset[2]</p>	<p>Predicción: 1, Clase: 0</p>  <p>Train_dataset[3]</p>
<p>Predicción: 0, Clase: 1</p>  <p>Train_dataset[4]</p>	<p>Predicción: 0, Clase: 1</p>  <p>Train_dataset[5]</p>	<p>Predicción: 0, Clase: 1</p>  <p>Train_dataset[6]</p>	<p>Predicción: 0, Clase: 1</p>  <p>Train_dataset[7]</p>
...
<p>Predicción: 1, Clase: 1</p>  <p>Train_dataset[4074]</p>	<p>Predicción: 1, Clase: 1</p>  <p>Train_dataset[4075]</p>	<p>Predicción: 1, Clase: 1</p>  <p>Train_dataset[4076]</p>	<p>Predicción: 1, Clase: 1</p>  <p>Train_dataset[4707]</p>

Figura 7: Vista Previa Imágenes posterior al entrenamiento