



BENEMÉRITA UNIVERSIDAD AUTÓNOMA DE PUEBLA

FACULTAD DE CIENCIAS DE LA ELECTRÓNICA  
MAESTRÍA EN INGENIERÍA ELECTRÓNICA,  
OPCIÓN INSTRUMENTACIÓN ELECTRÓNICA

Tesis para obtener el grado de:  
MAESTRO EN INGENIERÍA ELECTRÓNICA

---

Normalización y alineación automática de la forma de la región  
pulmonar integrada con selección de características discriminantes  
para detección de neumonía y COVID-19

---

Presenta:

Lic. Rafael Alejandro Cruz Ovando\*

Directores:

Dr. Salvador Eugenio Ayala Raggi

Dr. Aldrin Barreto Flores

# Índice general

Objetivos . . . . .	2
<b>1. Metodología</b>	<b>3</b>
1.1. Descripción General del Sistema . . . . .	3
1.1.1. Arquitectura del Sistema . . . . .	3
1.1.2. Flujo de Datos . . . . .	5
1.1.3. Justificación del Diseño Modular . . . . .	6
1.2. Conjunto de Datos y Preprocesamiento . . . . .	7
1.2.1. COVID-19 Radiography Database . . . . .	7
1.2.2. Anotación de Landmarks Anatómicos . . . . .	7
1.2.3. Preprocesamiento de Imágenes . . . . .	12
1.2.4. División del Conjunto de Datos . . . . .	14

# Objetivos

## Objetivo general

Desarrollar e implementar algoritmos de visión por computadora para la detección, alineación y normalización de la forma de la región pulmonar en imágenes radiográficas de tórax, utilizando además un método eficaz para la selección de características discriminantes, con el fin de mejorar la precisión en la detección automática de neumonía y COVID-19.

## Objetivos específicos

1. Diseñar, implementar y evaluar un método deformable de alineación y normalización que localice, segmente y ajuste automáticamente la región pulmonar en términos de forma, escala, posición y rotación.
2. Proponer un método de extracción y selección de características que maximicen la discriminación entre las clases.
3. Evaluar el rendimiento de diferentes clasificadores de aprendizaje supervisado para la técnica de alineación propuesta en la tesis: KNN, CNN, MLP.
4. Validar el clasificador desarrollado a través de medir la precisión, sensibilidad, especificidad y además de realizar pruebas de validación cruzada para caracterizar el algoritmo propuesto.
5. Contrastar los resultados de clasificación del objetivo anterior con resultados obtenidos por los mismos clasificadores pero sin realizar el proceso de alineación propuesto.
6. Publicación de resultados.

# Capítulo 1

## Metodología

Este capítulo presenta la metodología desarrollada para la normalización y alineación automática de radiografías de tórax, así como la clasificación de enfermedades pulmonares. Se describe el flujo de procesamiento completo del sistema, desde la adquisición de datos hasta la clasificación final, detallando cada componente y las decisiones de diseño tomadas.

### 1.1. Descripción General del Sistema

El desarrollo del sistema propuesto comprende dos fases: una fase de preparación, que incluye la anotación manual de landmarks anatómicos y el entrenamiento de los modelos, y una fase de operación, donde el sistema procesa nuevas radiografías de tórax. Durante la operación, las imágenes pasan por una secuencia de cuatro módulos: preprocesamiento, predicción de landmarks, normalización geométrica y clasificación. Los tres primeros módulos transforman la imagen de entrada a una representación geométricamente normalizada, mientras que el cuarto realiza la clasificación de diagnóstico. Este diseño modular permite evaluar la contribución de cada componente al rendimiento final del sistema.

La Figura 1.1 ilustra la relación entre ambas fases del sistema.

#### 1.1.1. Arquitectura del Sistema

La Figura 1.2 presenta el diagrama de bloques del flujo de operación. El sistema se compone de cuatro módulos que procesan secuencialmente las imágenes de entrada.

**Módulo 1: Preprocesamiento.** Las imágenes de entrada se someten a un proceso de mejora de contraste mediante el algoritmo CLAHE (*Contrast Limited Adaptive Histogram Equalization*) [1]. Este paso normaliza las variaciones de contraste inherentes a diferentes equipos de adquisición radiográfica. Posteriormente, las imágenes se redimensionan a  $224 \times 224$  píxeles para su procesamiento por la red neuronal.

**Módulo 2: Predicción de Landmarks.** Un modelo basado en ResNet-18 [2] con módulo

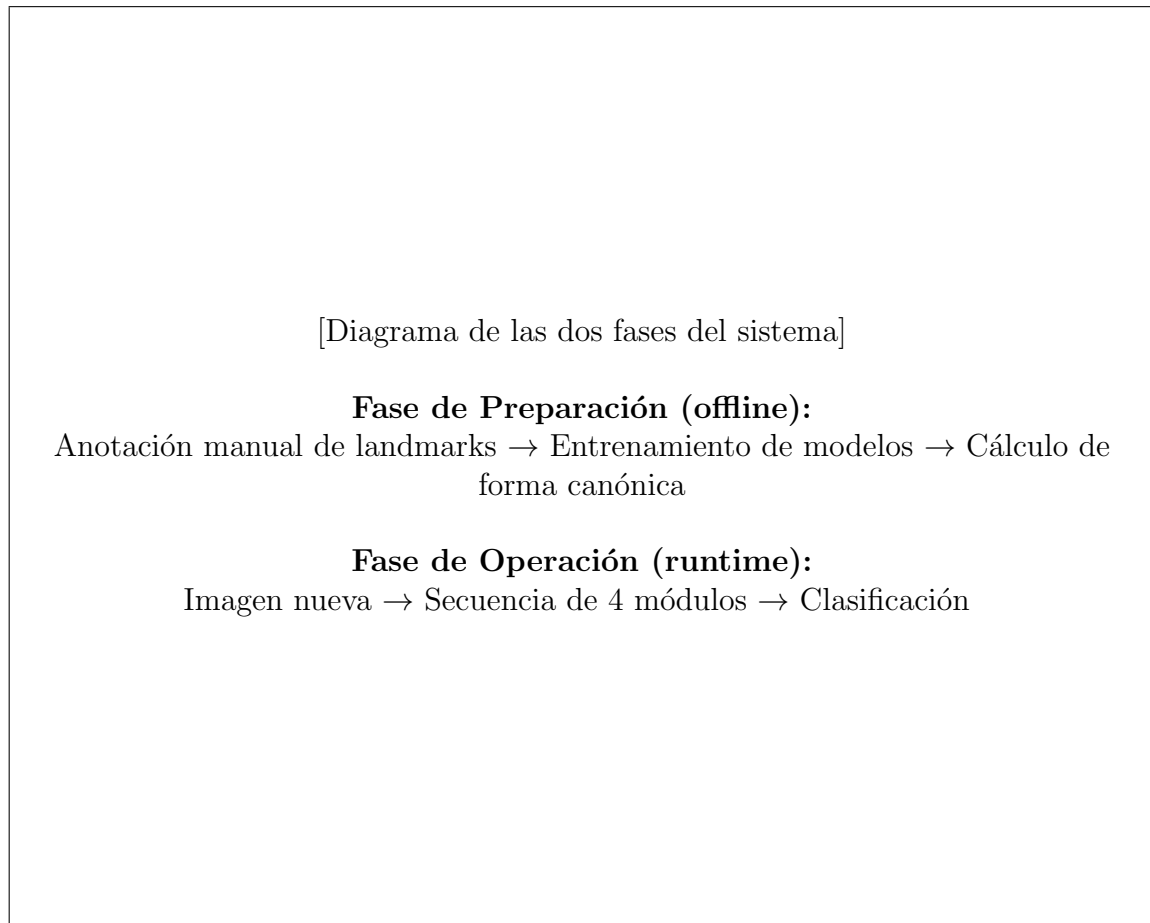


Figura 1.1: Estructura general del sistema en dos fases. La fase de preparación se ejecuta una única vez e incluye la anotación manual del conjunto de datos de entrenamiento, el entrenamiento de los modelos y el cálculo de la forma canónica mediante GPA. La fase de operación procesa cada imagen nueva a través de la secuencia de cuatro módulos.

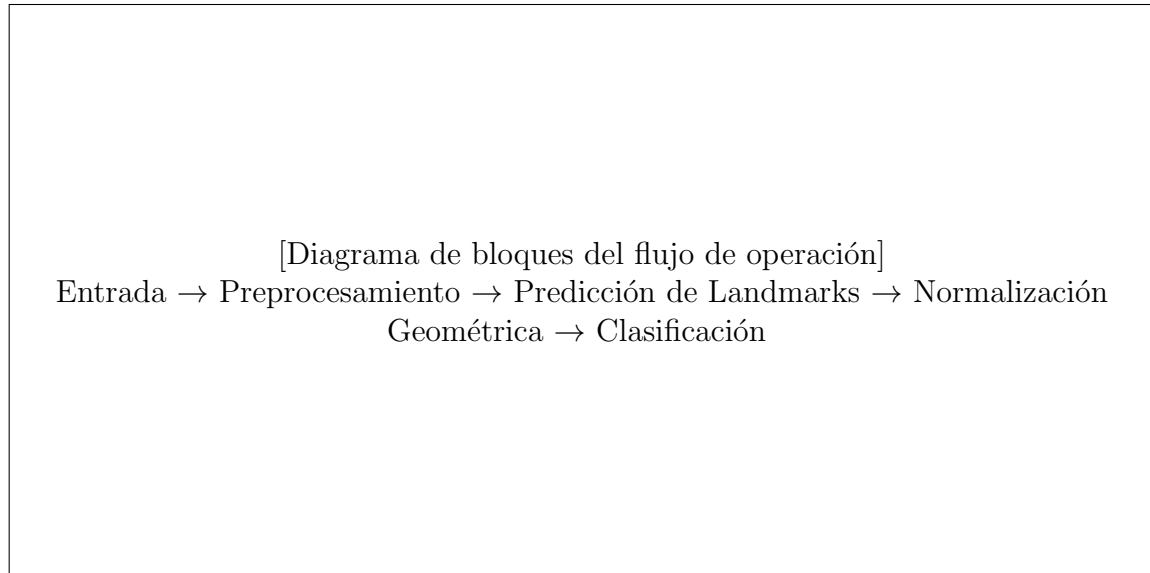


Figura 1.2: Flujo de operación del sistema. Las radiografías de tórax se procesan mediante cuatro módulos: preprocesamiento con CLAHE, predicción de 15 landmarks anatómicos, normalización geométrica mediante warping afín por partes, y clasificación en tres categorías (COVID-19, Normal, Neumonía Viral).

de Coordinate Attention [3] predice las coordenadas de 15 puntos anatómicos que definen el contorno de la región pulmonar. Estos landmarks fueron definidos manualmente durante la fase de anotación del conjunto de datos y representan puntos característicos de la silueta pulmonar bilateral.

**Módulo 3: Normalización Geométrica.** Utilizando los landmarks predichos, se aplica una transformación afín por partes (*piecewise affine warping*) que alinea cada imagen a una forma canónica previamente calculada mediante Análisis Procrustes Generalizado (GPA) [4]. Este proceso elimina variaciones geométricas entre pacientes, normalizando la posición, escala y orientación de la región pulmonar.

**Módulo 4: Clasificación.** Las imágenes normalizadas se procesan mediante una red neuronal convolucional para clasificarlas en una de tres categorías: COVID-19, Normal o Neumonía Viral. El módulo genera la predicción de clase junto con las probabilidades asociadas a cada categoría. Se evaluaron múltiples arquitecturas de clasificación, incluyendo ResNet-18, DenseNet-121 y EfficientNet-B0.

### 1.1.2. Flujo de Datos

El procesamiento de una imagen sigue el flujo ilustrado en la Tabla 1.1. Cada etapa transforma los datos de entrada en una representación apropiada para la siguiente etapa del proceso.

Cuadro 1.1: Flujo de datos a través del sistema.

Etapa	Entrada	Salida	Dimensiones
Preprocesamiento	Imagen RGB/Grayscale	Imagen normalizada	$224 \times 224 \times 3$
Predicción	Imagen normalizada	Coordenadas landmarks	$15 \times 2$
Warping	Imagen + landmarks	Imagen warped	$224 \times 224 \times 3$
Clasificación	Imagen warped	Vector probabilidades	3

### 1.1.3. Justificación del Diseño Modular

El diseño modular del sistema ofrece varias ventajas:

1. **Interpretabilidad:** Los landmarks predichos constituyen una representación intermedia que permite verificar visualmente la calidad del proceso de detección anatómica.
2. **Modularidad:** Cada componente puede entrenarse, evaluarse y mejorarse de forma independiente, facilitando el desarrollo iterativo del sistema.
3. **Selección implícita de características:** La normalización geométrica actúa como un mecanismo de selección de características a nivel de imagen, eliminando información no discriminante (artefactos, marcas hospitalarias, variaciones de pose) y preservando únicamente la región pulmonar relevante para la clasificación [5].
4. **Transferibilidad:** El modelo de landmarks puede reutilizarse para otras tareas de análisis pulmonar, mientras que el clasificador puede adaptarse a diferentes conjuntos de clases según los requerimientos de la aplicación.

El enfoque propuesto se fundamenta en la hipótesis de que la normalización geométrica mejora la robustez y capacidad de generalización del clasificador al reducir la variabilidad no relacionada con la patología. Esta hipótesis se evalúa experimentalmente en el Capítulo ??.

## 1.2. Conjunto de Datos y Preprocesamiento

Esta sección describe el conjunto de datos utilizado para el desarrollo y evaluación del sistema propuesto, así como los procesos de anotación y preprocesamiento aplicados a las imágenes.

### 1.2.1. COVID-19 Radiography Database

El presente trabajo utiliza el *COVID-19 Radiography Database*, un conjunto de datos públicamente disponible desarrollado por investigadores de Qatar University, University of Dhaka y colaboradores de Malasia y Pakistán [6, 7]. Este conjunto de datos ha sido ampliamente utilizado en la literatura para el desarrollo de sistemas de detección de COVID-19 basados en radiografías de tórax.

El conjunto de datos contiene imágenes de radiografías posteroanterior (PA) de tórax organizadas en tres categorías diagnósticas:

- **COVID-19:** Radiografías de pacientes con diagnóstico confirmado de COVID-19 mediante prueba RT-PCR.
- **Normal:** Radiografías de pacientes sin patología pulmonar aparente.
- **Neumonía Viral:** Radiografías de pacientes con neumonía viral de etiología distinta a SARS-CoV-2.

La Tabla 1.2 presenta la distribución de imágenes por categoría utilizada en este trabajo.

Cuadro 1.2: Distribución del conjunto de datos por categoría diagnóstica.

Categoría	Imágenes	Porcentaje	Fuente
COVID-19	3,616	23.9 %	BIMCV, Eurorad, GitHub
Normal	10,192	67.2 %	RSNA, CheXpert
Neumonía Viral	1,345	8.9 %	Chest X-Ray Images
<b>Total</b>	<b>15,153</b>	<b>100 %</b>	—

Las imágenes originales tienen un tamaño de  $299 \times 299$  píxeles en formato PNG. El conjunto de datos presenta un desbalance de clases natural, con predominancia de imágenes normales, lo cual refleja la distribución típica en escenarios clínicos reales.

### 1.2.2. Anotación de Landmarks Anatómicos

Para el entrenamiento del modelo de predicción de landmarks, se realizó la anotación manual de 15 puntos característicos en un subconjunto del conjunto de datos. Estos landmarks



definen el contorno de la región pulmonar bilateral y fueron seleccionados para permitir una triangulación robusta en el proceso de normalización geométrica.

### Definición de los 15 Landmarks

Los landmarks no corresponden a estructuras anatómicas específicas, sino que representan puntos de control sobre la silueta pulmonar diseñados para capturar la forma global del contorno. La Figura 1.3 ilustra la ubicación de cada landmark.

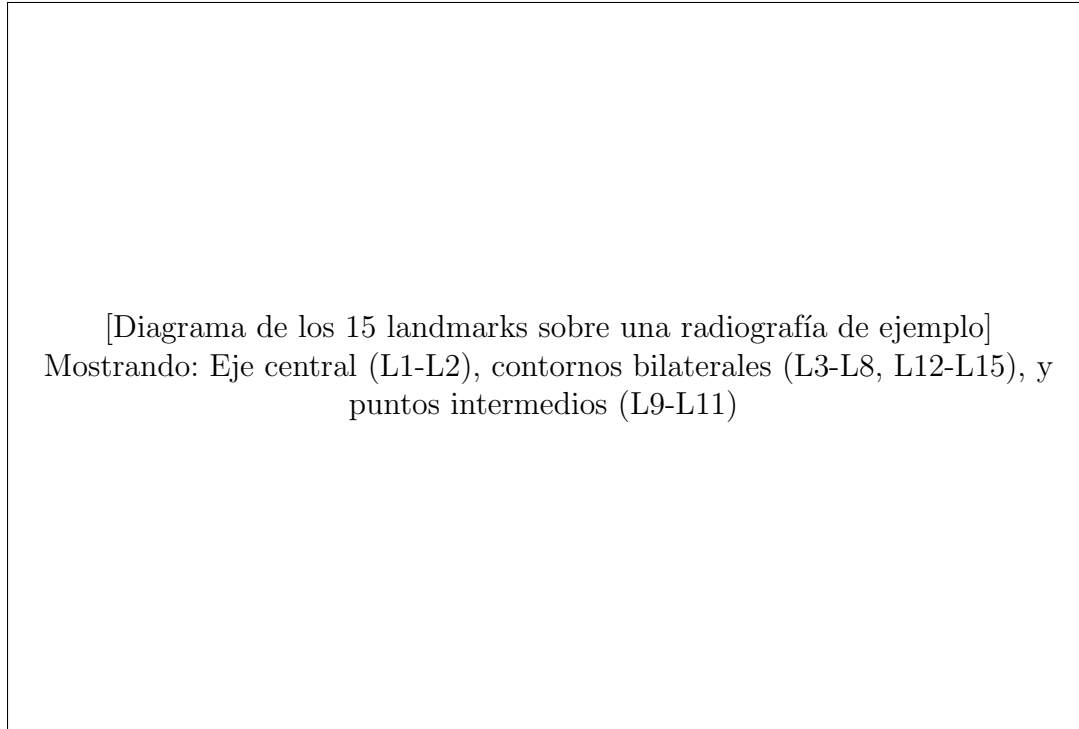


Figura 1.3: Ubicación de los 15 landmarks que definen el contorno pulmonar. Los puntos L1 y L2 definen el eje vertical central. Los landmarks L3-L8 delimitan los contornos laterales, mientras que L9-L11 dividen el eje central en cuatro segmentos iguales. Los pares (L12,L13) y (L14,L15) corresponden a las esquinas superior e inferior respectivamente.

La estructura geométrica de los landmarks se organiza de la siguiente manera:

1. **Eje central vertical:** Los landmarks L1 (superior) y L2 (inferior) definen la línea media de la silueta pulmonar. Los puntos L9, L10 y L11 dividen este eje en cuatro segmentos de igual longitud, correspondiendo a las posiciones relativas  $t = 0,25$ ,  $t = 0,50$  y  $t = 0,75$  respectivamente.
2. **Contorno pulmonar izquierdo:** Los landmarks L12, L3, L5, L7 y L14 trazan el borde lateral izquierdo de la silueta, desde la región superior hasta la inferior.

3. **Contorno pulmonar derecho:** De manera simétrica, los landmarks L13, L4, L6, L8 y L15 definen el borde lateral derecho.
4. **Pares simétricos:** Existen cinco pares de landmarks bilateralmente simétricos: (L3, L4), (L5, L6), (L7, L8), (L12, L13) y (L14, L15). Esta simetría estructural facilita la transformación de coordenadas durante el aumento de datos (flip horizontal).

La Tabla 1.3 resume la ubicación y función de cada landmark.

Cuadro 1.3: Descripción de los 15 landmarks anatómicos.

ID	Ubicación	Posición $t$	Par simétrico
L1	Ápex (eje central superior)	0.00	—
L2	Base (eje central inferior)	1.00	—
L3	Contorno izquierdo superior	0.25	L4
L4	Contorno derecho superior	0.25	L3
L5	Contorno izquierdo medio	0.50	L6
L6	Contorno derecho medio	0.50	L5
L7	Contorno izquierdo inferior	0.75	L8
L8	Contorno derecho inferior	0.75	L7
L9	Eje central (cuarto superior)	0.25	—
L10	Eje central (punto medio)	0.50	—
L11	Eje central (cuarto inferior)	0.75	—
L12	Esquina superior izquierda	0.00	L13
L13	Esquina superior derecha	0.00	L12
L14	Esquina inferior izquierda	1.00	L15
L15	Esquina inferior derecha	1.00	L14

## Proceso de Anotación

Para realizar la anotación de landmarks se desarrolló una herramienta gráfica interactiva basada en OpenCV que facilita el proceso mediante un algoritmo semi-automático. La anotación se realizó sobre un subconjunto de 957 imágenes seleccionadas del conjunto de datos completo, asegurando representatividad de las tres categorías diagnósticas.

**Herramienta de Anotación** La herramienta desarrollada implementa un proceso de anotación en dos fases que reduce significativamente el tiempo requerido respecto a la marcación individual de cada punto. La Figura 1.4 ilustra la interfaz de la herramienta.

**Fase 1: Generación Automática** El proceso inicia con tres interacciones del operador que definen la geometría base:

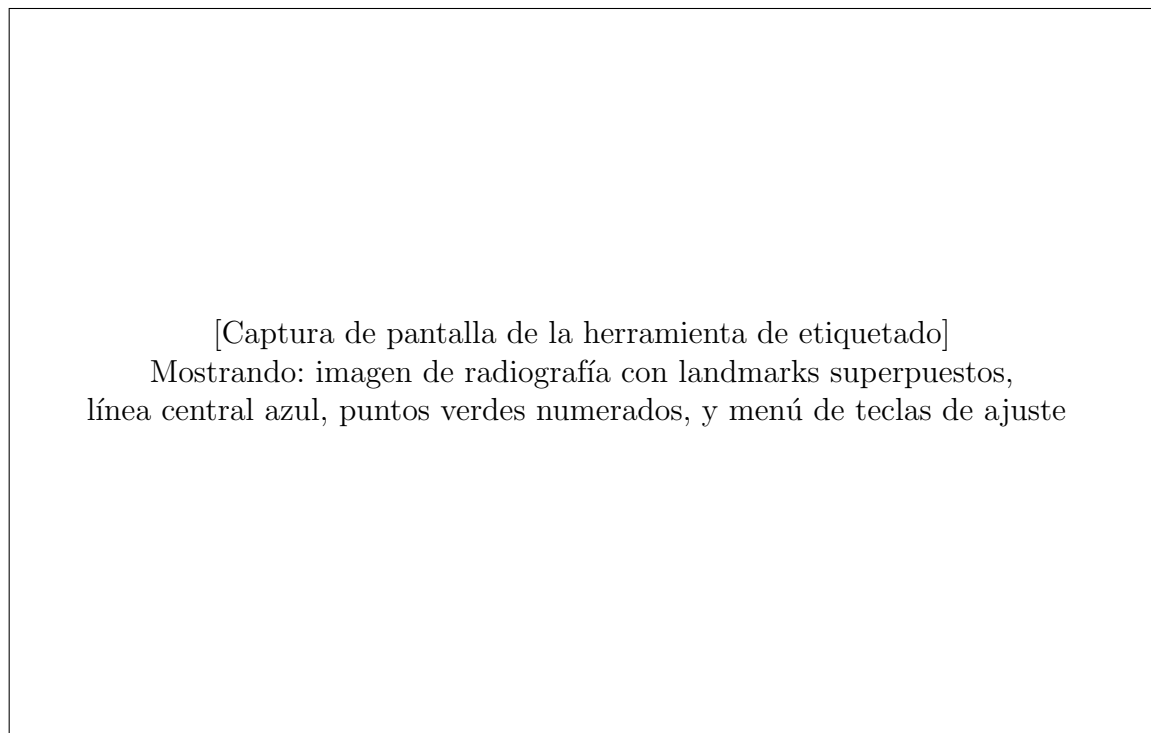


Figura 1.4: Interfaz de la herramienta de anotación de landmarks. La ventana principal muestra la radiografía con una línea vertical central de referencia. Los landmarks se visualizan como puntos verdes conectados por líneas rojas que definen el contorno pulmonar.

1. **Primer click (L1):** Define el ápex superior de la silueta pulmonar.
2. **Segundo click (L2):** Define la base inferior, estableciendo el eje central del contorno.
3. **Tercer click:** Confirma la selección y activa el algoritmo de generación automática.

El algoritmo de generación automática calcula los 13 landmarks restantes (L3-L15) mediante el siguiente procedimiento:

1. Calcula la línea central entre L1 y L2, determinando su pendiente.
2. Divide el eje central en cuatro segmentos iguales, ubicando los puntos intermedios L9, L10 y L11 en las posiciones  $t = 0,25$ ,  $t = 0,50$  y  $t = 0,75$  respectivamente.
3. Genera líneas perpendiculares al eje central en cada punto de división.
4. Ubica los landmarks laterales (L3-L8, L12-L15) sobre estas perpendiculares a distancias predefinidas del eje central.

**Fase 2: Ajuste Manual** Los landmarks generados automáticamente proporcionan una aproximación inicial que raramente coincide exactamente con el contorno pulmonar visible. La herramienta permite ajustar cada landmark horizontalmente mediante atajos de teclado, manteniendo la coherencia geométrica al desplazar los puntos a lo largo de sus respectivas líneas perpendiculares.

El ajuste se realiza hasta que cada landmark coincida visualmente con el borde de la silueta pulmonar en la imagen.

**Criterios de Anotación** El proceso de anotación siguió las siguientes directrices para garantizar consistencia:

1. Se colocó cada landmark sobre el borde perceptible de la silueta pulmonar, no sobre estructuras anatómicas internas.
2. En casos de ambigüedad por baja calidad de imagen o superposición de estructuras, se priorizó la consistencia visual sobre la precisión anatómica.
3. Se verificó visualmente que los pares simétricos (L3-L4, L5-L6, etc.) mantuvieran una distribución razonable respecto al eje central.
4. Las coordenadas se registraron en píxeles respecto a la imagen original de  $299 \times 299$  píxeles.

Las anotaciones se almacenaron en formato CSV con la estructura mostrada en la Tabla 1.4.

La distribución del subconjunto anotado por categoría se presenta en la Tabla 1.5.

Cuadro 1.4: Formato del archivo CSV de coordenadas de landmarks.

Campo	Descripción
índice	Identificador numérico de la imagen
L1_x, L1_y, ..., L15_x, L15_y	Coordenadas $(x, y)$ de cada landmark en píxeles
image_name	Nombre del archivo de imagen (categoría-ID)

Cuadro 1.5: Distribución del subconjunto anotado con landmarks.

Categoría	Imágenes anotadas	Porcentaje
COVID-19	306	32.0 %
Normal	468	48.9 %
Neumonía Viral	183	19.1 %
<b>Total</b>	<b>957</b>	<b>100 %</b>

### 1.2.3. Preprocesamiento de Imágenes

Las imágenes radiográficas requieren preprocesamiento para mitigar las variaciones introducidas por distintos equipos de adquisición y diversas condiciones de exposición. El proceso implementado consta de tres etapas: mejora de contraste, redimensionamiento y normalización.

#### Mejora de Contraste mediante CLAHE

Se aplica el algoritmo *Contrast Limited Adaptive Histogram Equalization* (CLAHE) [8] para mejorar el contraste local de las imágenes. A diferencia de la ecualización de histograma global, CLAHE opera sobre regiones locales (tiles) y limita la amplificación de contraste para evitar el realce excesivo de ruido.

Los parámetros utilizados fueron determinados experimentalmente:

- **Clip limit:** 2,0 — Controla el límite máximo de amplificación de contraste. Valores mayores producen mayor contraste pero pueden amplificar ruido.
- **Tile size:**  $4 \times 4$  — Tamaño de las regiones para ecualización local. Un valor menor produce una adaptación más fina pero aumenta el tiempo de cómputo.

La Figura 1.5 muestra el efecto del preprocesamiento CLAHE sobre una radiografía de ejemplo.



[Comparación lado a lado: imagen original vs. imagen con CLAHE aplicado]

Figura 1.5: Efecto del preprocesamiento CLAHE. (a) Imagen original con bajo contraste en la región pulmonar. (b) Imagen después de aplicar CLAHE con clip limit = 2,0 y tile size = 4, mostrando mejor definición de estructuras pulmonares.

## Redimensionamiento

Las imágenes se redimensionan de su tamaño original ( $299 \times 299$  píxeles) a  $224 \times 224$  píxeles mediante interpolación bilineal. Este tamaño corresponde a la entrada estándar de las arquitecturas de redes neuronales preentrenadas en ImageNet [9].

## Normalización

Para el modelo de predicción de landmarks basado en ResNet-18 preentrenado, se aplica normalización utilizando las estadísticas del conjunto de datos ImageNet:

$$\hat{x}_c = \frac{x_c - \mu_c}{\sigma_c} \quad (1.1)$$

donde  $x_c$  es el valor del canal  $c$ ,  $\mu_c$  es la media y  $\sigma_c$  la desviación estándar para cada canal RGB:

$$\mu = (0,485, 0,456, 0,406) \quad (1.2)$$

$$\sigma = (0,229, 0,224, 0,225) \quad (1.3)$$

Las coordenadas de los landmarks se normalizan al rango  $[0, 1]$  dividiendo entre el tamaño de la imagen (224 píxeles), facilitando el entrenamiento del modelo de regresión.

### 1.2.4. División del Conjunto de Datos

El conjunto de datos se divide en tres subconjuntos mutuamente excluyentes para entrenamiento, validación y prueba. La división se realiza de manera estratificada por categoría diagnóstica para mantener las proporciones de clases en cada subconjunto.

- **Entrenamiento (75 %):** Utilizado para optimizar los parámetros del modelo.
- **Validación (15 %):** Utilizado para selección de hiperparámetros y criterio de parada temprana.
- **Prueba (10 %):** Reservado exclusivamente para la evaluación final del modelo.

La Tabla 1.6 presenta la distribución resultante para el conjunto de datos de clasificación (15,153 imágenes).

Cuadro 1.6: División del conjunto de datos en subconjuntos de entrenamiento, validación y prueba.

Conjunto	COVID-19	Normal	Viral	Total
Entrenamiento (75 %)	2,712	7,644	1,008	11,364
Validación (15 %)	543	1,529	202	2,274
Prueba (10 %)	361	1,019	135	1,515
<b>Total</b>	<b>3,616</b>	<b>10,192</b>	<b>1,345</b>	<b>15,153</b>

Para garantizar la reproducibilidad de los experimentos, se utiliza una semilla aleatoria fija ( $seed = 42$ ) en todas las operaciones que involucran aleatorización, incluyendo la división del conjunto de datos, inicialización de pesos del modelo y muestreo durante el entrenamiento.

El subconjunto con anotaciones de landmarks (957 imágenes) sigue la misma proporción de división:

- Entrenamiento: 717 imágenes (75 %)
- Validación: 144 imágenes (15 %)
- Prueba: 96 imágenes (10 %)

# Bibliografía

- [1] K. Zuiderveld, “Contrast limited adaptive histogram equalization,” in *Graphics gems IV*. Academic Press Professional, 1994, pp. 474–485.
- [2] K. He, X. Zhang, S. Ren, and J. Sun, “Deep residual learning for image recognition,” in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2016, pp. 770–778.
- [3] Q. Hou, D. Zhou, and J. Feng, “Coordinate attention for efficient mobile network design,” in *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, 2021, pp. 13 713–13 722.
- [4] J. C. Gower, “Generalized procrustes analysis,” *Psychometrika*, vol. 40, no. 1, pp. 33–51, 1975.
- [5] M. Jaderberg, K. Simonyan, A. Zisserman, and K. Kavukcuoglu, “Spatial transformer networks,” in *Advances in neural information processing systems*, vol. 28, 2015.
- [6] M. E. Chowdhury, T. Rahman, A. Khandakar, R. Mazhar, M. A. Kadir, Z. B. Mahbub, K. R. Islam, M. S. Khan, A. Iqbal, N. Al Emadi *et al.*, “Can ai help in screening viral and covid-19 pneumonia?” *IEEE Access*, vol. 8, pp. 132 665–132 676, 2020.
- [7] T. Rahman, A. Khandakar, Y. Qiblawey, A. Tahir, S. Kiranyaz, S. B. A. Kashem, M. T. Islam, S. Al Maadeed, S. M. Zughaier, M. S. Khan *et al.*, “Exploring the effect of image enhancement techniques on covid-19 detection using chest x-ray images,” *Computers in Biology and Medicine*, vol. 132, p. 104319, 2021.
- [8] S. M. Pizer, E. P. Amburn, J. D. Austin, R. Cromartie, A. Geselowitz, T. Greer, B. ter Haar Romeny, J. B. Zimmerman, and K. Zuiderveld, “Adaptive histogram equalization and its variations,” *Computer Vision, Graphics, and Image Processing*, vol. 39, no. 3, pp. 355–368, 1987.
- [9] J. Deng, W. Dong, R. Socher, L.-J. Li, K. Li, and L. Fei-Fei, “Imagenet: A large-scale hierarchical image database,” in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2009, pp. 248–255.