

# Normalización geométrica de la región pulmonar mediante predicción de landmarks para clasificación automática de neumonía y COVID-19

Rafael Alejandro Cruz Ovando<sup>1</sup>, Salvador Eugenio Ayala Raggi<sup>1</sup>, and Aldrin Barreto Flores<sup>1</sup>

Benemérita Universidad Autónoma de Puebla, Facultad de Ciencias de la  
Electrónica, Puebla, México  
co223470443@alm.buap.mx

**Resumen** Este artículo presenta un sistema de clasificación automática de enfermedades pulmonares en radiografías de tórax basado en normalización geométrica. El método propuesto consta de tres etapas principales. **Primera etapa:** se entrenan redes neuronales convolucionales ResNet-18 con mecanismo de atención por coordenadas para predecir 15 puntos de referencia (landmarks) que definen el contorno pulmonar, logrando un error medio de 3.61 píxeles mediante un ensemble de cuatro modelos con aumento de datos en tiempo de prueba. **Segunda etapa:** las coordenadas predichas se utilizan para normalizar geométricamente cada imagen mediante Análisis Procrustes Generalizado (para obtener una forma pulmonar estándar), triangulación de Delaunay (para construir una malla de deformación) y transformación afín por partes (warping), eliminando variaciones de posición, escala y orientación. **Tercera etapa:** las imágenes normalizadas se clasifican en tres categorías (COVID-19, Neumonía Viral y Normal) mediante un clasificador ResNet-18 con aprendizaje por transferencia y preprocesamiento SAHS. El sistema fue evaluado sobre la base de datos COVID-19 Radiography Database con 15,153 imágenes, alcanzando una exactitud del 98.10 % y F1-Macro del 97.17 %. La comparación controlada entre configuraciones demuestra que la normalización geométrica permite obtener rendimiento comparable a las imágenes originales (diferencia de 0.58 puntos porcentuales), mientras que el simple recorte de bordes sin normalización produce resultados inferiores. Estos resultados evidencian que el sistema propuesto aprende características patológicas genuinas de la región pulmonar.

**Keywords:** Normalización geométrica · Predicción de landmarks · COVID-19 · Clasificación de imágenes médicas · Aprendizaje profundo

## 1. Introducción

La neumonía representa una de las principales causas de morbilidad y mortalidad a nivel mundial, siendo especialmente crítica durante la pandemia de

COVID-19. El diagnóstico oportuno y preciso de esta enfermedad es fundamental para el tratamiento adecuado de los pacientes. Las radiografías de tórax constituyen una herramienta diagnóstica de primera línea debido a su amplia disponibilidad, bajo costo y rapidez de adquisición [18]. Sin embargo, la interpretación manual de estas imágenes requiere experiencia especializada y está sujeta a variabilidad inter-observador, lo que motiva el desarrollo de métodos automáticos de clasificación.

En años recientes, los métodos basados en aprendizaje profundo han demostrado gran potencial para la detección automática de neumonía en radiografías [15,16,2]. Estos enfoques utilizan redes neuronales convolucionales (CNNs) para extraer características discriminantes directamente de las imágenes. Sin embargo, su rendimiento puede verse afectado por variaciones en la adquisición de las imágenes, incluyendo diferencias de posición, escala y orientación del paciente, así como artefactos propios de cada equipo radiográfico. Este fenómeno, conocido como *domain shift* [20], puede comprometer la capacidad de generalización de los modelos.

### 1.1. Trabajo Relacionado

Diversos trabajos han abordado el problema de la variabilidad en imágenes médicas mediante técnicas de normalización y alineación. Ayala-Raggi et al. [1] demostraron que la normalización de radiografías de tórax, combinada con selección de características discriminantes, mejora significativamente el reconocimiento automático de COVID-19. En su estudio presentado en ACPR 2023, los autores mostraron que eliminar variaciones no patológicas permite a los clasificadores enfocarse en las características intrínsecas de la enfermedad, resultando en mayor exactitud diagnóstica.

De manera complementaria, Picazo-Castillo et al. [13] realizaron un estudio comparativo de diferentes representaciones de imágenes pulmonares para el reconocimiento automático de neumonía. Sus resultados evidencian que la forma en que se presenta la información visual al clasificador tiene un impacto significativo en el rendimiento, destacando la importancia de preprocesar adecuadamente las imágenes para facilitar el aprendizaje de patrones patológicos.

Otros trabajos han explorado el uso de transformaciones geométricas en imágenes médicas. Las Spatial Transformer Networks [9] permiten aprender transformaciones espaciales de manera end-to-end, pero requieren grandes cantidades de datos para el entrenamiento. En contraste, los métodos basados en landmarks anatómicos ofrecen una alternativa más interpretable y pueden funcionar con conjuntos de datos más pequeños, ya que incorporan conocimiento a priori sobre la anatomía.

### 1.2. Propuesta del Presente Trabajo

En este trabajo proponemos un método de normalización geométrica que va más allá de las transformaciones rígidas tradicionales (rotación, traslación, escalado). El sistema propuesto incorpora deformaciones locales basadas en la

anatomía pulmonar para transformar cada radiografía a una forma estándar, eliminando no solo variaciones de pose global sino también diferencias en la morfología del tórax entre pacientes.

El método se fundamenta en tres componentes principales. Primero, se emplean redes neuronales convolucionales para predecir automáticamente las coordenadas de 15 puntos de referencia que definen el contorno pulmonar bilateral. Segundo, se utiliza el Análisis Procrustes Generalizado [6] sobre un conjunto de landmarks anotados manualmente para determinar una forma estándar pulmonar representativa del dataset. Tercero, se aplica triangulación de Delaunay [3] para construir una malla de triángulos sobre los landmarks, y se realiza una transformación afin por partes (piecewise affine warping) [17] que deforma cada imagen para hacer coincidir sus landmarks predichos con la forma estándar. Las imágenes normalizadas resultantes se utilizan para entrenar un clasificador ResNet-18 [7] con aprendizaje por transferencia.

Este enfoque permite obtener una representación estandarizada de la región pulmonar donde la variabilidad no patológica (debida a diferencias anatómicas, de posición o de adquisición) ha sido minimizada, permitiendo que el clasificador se enfoque en aprender las características intrínsecas de cada patología.

### 1.3. Organización del Artículo

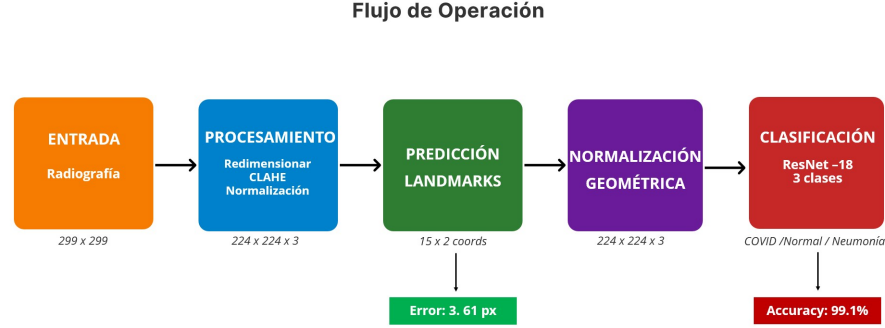
El resto del artículo se organiza como sigue. La Sección 2 describe detalladamente la metodología propuesta, incluyendo la arquitectura de las redes, el proceso de normalización geométrica y el entrenamiento del clasificador. La Sección 3 presenta los resultados experimentales, abordando la precisión de localización de landmarks, el efecto de la normalización en la clasificación, y la robustez del sistema. Finalmente, la Sección 4 presenta las conclusiones y direcciones de trabajo futuro.

## 2. Metodología

### 2.1. Vista General del Sistema

Desde una perspectiva de alto nivel, el sistema propuesto puede entenderse como una “caja negra” que recibe una radiografía de tórax como entrada y produce un diagnóstico (Normal, COVID-19 o Neumonía Viral) como salida. Internamente, el sistema ejecuta tres etapas de procesamiento: (1) predicción de puntos de referencia anatómicos, (2) normalización geométrica de la imagen, y (3) clasificación de la imagen normalizada. Esta arquitectura modular permite que cada componente sea entrenado y evaluado de manera independiente.

La Figura 1 ilustra el flujo de procesamiento completo. Una imagen de entrada pasa primero por el modelo de predicción de landmarks, que genera 15 pares de coordenadas  $(x, y)$  describiendo el contorno pulmonar. Estas coordenadas, junto con la forma estándar pre-calculada, se utilizan para realizar la transformación geométrica. La imagen resultante, donde la región pulmonar ha sido normalizada a una forma estándar, se procesa finalmente por el clasificador que emite el diagnóstico.



**Figura 1.** Flujo de operación del sistema propuesto. De izquierda a derecha: (1) entrada de radiografía ( $299 \times 299$ ), (2) preprocesamiento con redimensionado, CLAHE y normalización ( $224 \times 224 \times 3$ ), (3) predicción de 15 pares de coordenadas con error de 3.61 px, (4) normalización geométrica mediante warping ( $224 \times 224 \times 3$ ), y (5) clasificación con ResNet-18 en tres categorías (COVID-19, Normal, Neumonía Viral).

## 2.2. Conjunto de Datos

Se utilizó la base de datos pública **COVID-19 Radiography Database** [2,14], que contiene 15,153 radiografías posteroanteriores organizadas en tres categorías: COVID-19 (3,616 imágenes), Normal (10,192 imágenes) y Neumonía Viral (1,345 imágenes). Las imágenes originales ( $299 \times 299$  píxeles) se redimensionaron a  $224 \times 224$  píxeles, formato estándar para modelos preentrenados en ImageNet.

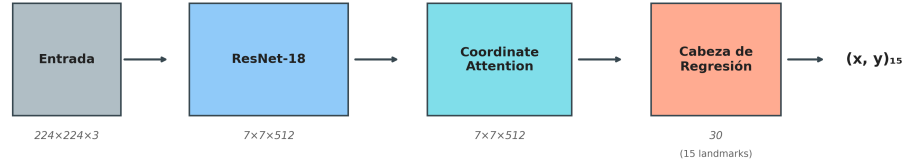
Para el entrenamiento supervisado del modelo de predicción de landmarks, se anotaron manualmente 15 puntos de referencia en el contorno pulmonar de 957 imágenes. Los landmarks definen puntos de control sobre la silueta pulmonar bilateral, organizados en un eje central vertical (5 puntos) y cinco pares simétricos izquierdo/derecho (10 puntos). El conjunto de datos se dividió de manera estratificada en entrenamiento (75 %), validación (15 %) y prueba (10 %).

## 2.3. Arquitectura del Modelo de Predicción de Landmarks

El modelo de predicción se basa en una arquitectura ResNet-18 [7] preentrenada en ImageNet [4]. Este enfoque de aprendizaje por transferencia aprovecha las características visuales aprendidas de millones de imágenes (bordes, texturas, patrones) como base para la tarea específica de localización de landmarks, lo cual es particularmente beneficioso dado el tamaño limitado del conjunto de datos anotado.

La arquitectura se estructura en tres componentes principales (Figura 2):

- **Backbone:** Las capas convolucionales de ResNet-18 procesan la imagen de entrada ( $224 \times 224 \times 3$ ) produciendo un mapa de características de  $7 \times 7 \times 512$ .
- **Módulo Coordinate Attention:** Insertado entre el backbone y la cabeza de regresión, este mecanismo [8] captura dependencias posicionales a lo largo de ambas dimensiones espaciales, crucial para tareas de localización precisa.
- **Cabeza de regresión:** Tres capas completamente conectadas ( $512 \rightarrow 512 \rightarrow 768 \rightarrow 30$ ) con Group Normalization [19] y dropout escalonado. La salida son 30 valores normalizados en  $[0,1]$ , correspondientes a las coordenadas  $(x, y)$  de los 15 landmarks.



**Figura 2.** Arquitectura del modelo de predicción de landmarks. La imagen de entrada ( $224 \times 224 \times 3$ ) pasa por el backbone ResNet-18 que produce un mapa de características ( $7 \times 7 \times 512$ ), luego por el módulo Coordinate Attention que mantiene las dimensiones, y finalmente por la cabeza de regresión que produce 30 valores correspondientes a las coordenadas  $(x, y)$  de los 15 landmarks.

## 2.4. Entrenamiento del Modelo de Landmarks

Se empleó Wing Loss [5] como función de pérdida, diseñada específicamente para localización de landmarks por su comportamiento adaptativo: logarítmico para errores pequeños (refinamiento fino) y lineal para errores grandes (estabilidad):

$$\mathcal{L}_{\text{wing}}(e) = \begin{cases} \omega \ln \left( 1 + \frac{|e|}{\epsilon} \right), & |e| < \omega \\ |e| - C, & |e| \geq \omega \end{cases} \quad (1)$$

con parámetros  $\omega = 10$  píxeles,  $\epsilon = 2$  píxeles, y  $C = \omega - \omega \ln(1 + \omega/\epsilon)$ .

El entrenamiento se organizó en dos fases para un ajuste estable. En la primera fase (15 épocas), solo se entrenó la cabeza de regresión con tasa de aprendizaje  $1 \times 10^{-3}$ , manteniendo congelados el backbone y Coordinate Attention. En la

segunda fase (100 épocas máximo), todas las capas se entrenaron con tasas diferenciadas:  $2 \times 10^{-5}$  para backbone y Coordinate Attention, y  $2 \times 10^{-4}$  para la cabeza.

Para mejorar la precisión, se implementó un **ensemble** de cuatro modelos idénticos entrenados con diferentes semillas aleatorias. Durante la inferencia, se aplicó **Test-Time Augmentation (TTA)** mediante volteo horizontal, promediando las predicciones originales y reflejadas.

## 2.5. Análisis Procrustes Generalizado

Una vez entrenado el modelo de predicción, el siguiente paso es determinar una forma estándar pulmonar que sirva como referencia para la normalización. Para esto se aplicó el Análisis Procrustes Generalizado (GPA) [6] sobre las 957 configuraciones de landmarks anotadas manualmente.

El GPA es un algoritmo iterativo que elimina variaciones de traslación, escala y rotación mediante tres operaciones por configuración: (1) centrado al origen, (2) escalado a norma unitaria, y (3) rotación óptima calculada por Descomposición en Valores Singulares (SVD). El proceso se repite hasta convergencia (tolerancia  $\tau = 10^{-8}$ ), produciendo una forma estándar que representa la morfología pulmonar promedio del dataset (Figura 3).

## 2.6. Triangulación de Delaunay y Warping

Sobre los 15 landmarks de la forma estándar, se calculó una triangulación de Delaunay [3], resultando en una malla de 16 triángulos que cubre la región pulmonar (Figura 4). Esta malla define la correspondencia entre puntos de la imagen original y la forma estándar.

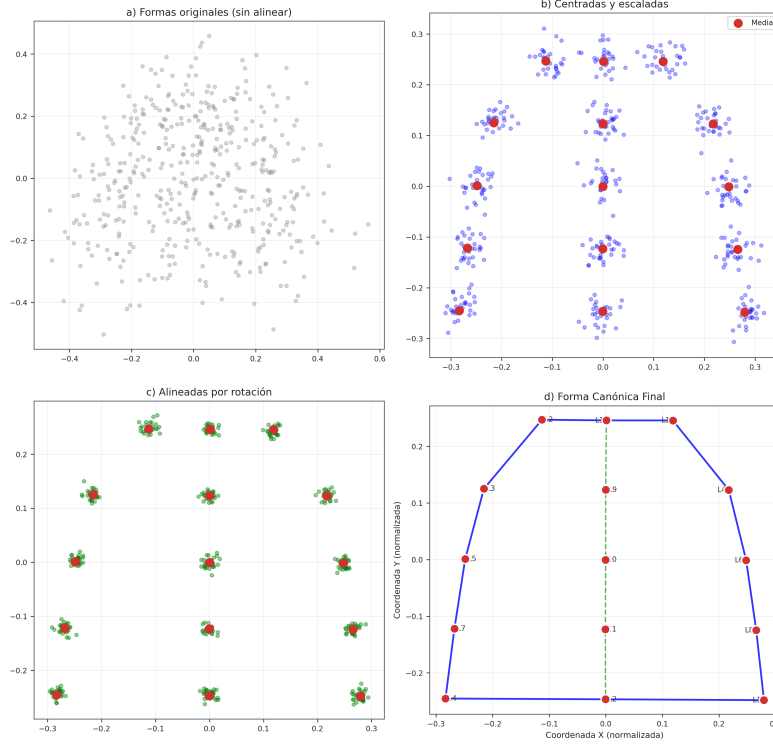
Para cada par de triángulos correspondientes (imagen original  $\leftrightarrow$  forma estándar), se calculó una transformación afín única [17] que mapea los tres vértices del triángulo original a sus posiciones en la forma estándar. La transformación preserva continuidad en los bordes triangulares y utiliza interpolación bilineal para suavizar la imagen resultante.

Un parámetro clave, `margin_scale` = 1.05, controla una leve expansión radial desde el centroide de los landmarks, asegurando que se incluya contexto anatómico periférico relevante. Este proceso produce imágenes donde la región pulmonar ocupa aproximadamente el 47% del área total (Figura 5).

## 2.7. Generación del Dataset Normalizado

Una vez entrenado el modelo de landmarks y calculada la forma estándar, se procedió a generar el dataset completo de imágenes normalizadas. El modelo ensemble se utilizó para inferir los landmarks en las 15,153 imágenes del dataset, y cada imagen se transformó mediante el proceso de warping descrito. Este dataset normalizado constituye la entrada para el entrenamiento del clasificador final.

Análisis Procrustes Generalizado (GPA)

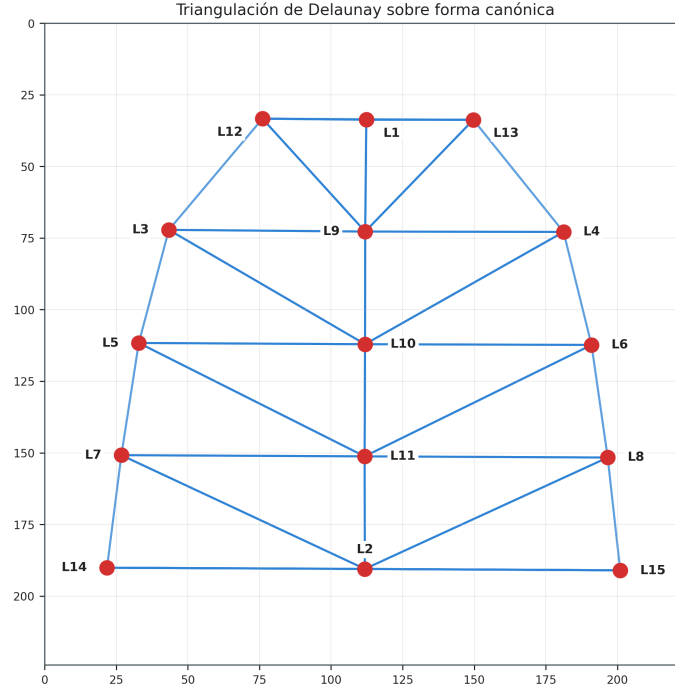


**Figura 3.** Proceso de Análisis Procrustes Generalizado en cuatro etapas. (a) Formas originales sin alinear mostrando alta dispersión. (b) Configuraciones centradas y escaladas con puntos azules agrupados alrededor de la media (rojo). (c) Configuraciones alineadas por rotación con mayor concentración. (d) Forma estándar final con los 15 landmarks etiquetados (L1–L15) definiendo el contorno pulmonar.

## 2.8. Clasificación

Para la clasificación multiclase se empleó una arquitectura ResNet-18 preentrenada en ImageNet, reemplazando su capa final por una nueva capa completamente conectada con 3 salidas (COVID-19, Normal, Neumonía Viral). Previo a la clasificación, las imágenes normalizadas se procesaron con **SAHS (Statistical Asymmetrical Histogram Stretching)**, un método de mejora de contraste diseñado específicamente para radiografías.

El dataset presenta desbalance significativo de clases (67% Normal, 24% COVID-19, 9% Neumonía Viral). Para mitigar el sesgo hacia la clase mayoritaria, se utilizaron pesos inversamente proporcionales a la frecuencia de cada clase en la función de pérdida Cross-Entropy.



**Figura 4.** Triangulación de Delaunay sobre la forma estándar. Los 15 landmarks (L1–L15, puntos rojos) están conectados por aristas azules formando una malla de 16 triángulos. El eje central (L1, L9, L10, L11, L2) divide la silueta pulmonar, mientras los pares simétricos (L3/L4, L5/L6, L7/L8, L12/L13, L14/L15) definen el contorno bilateral.

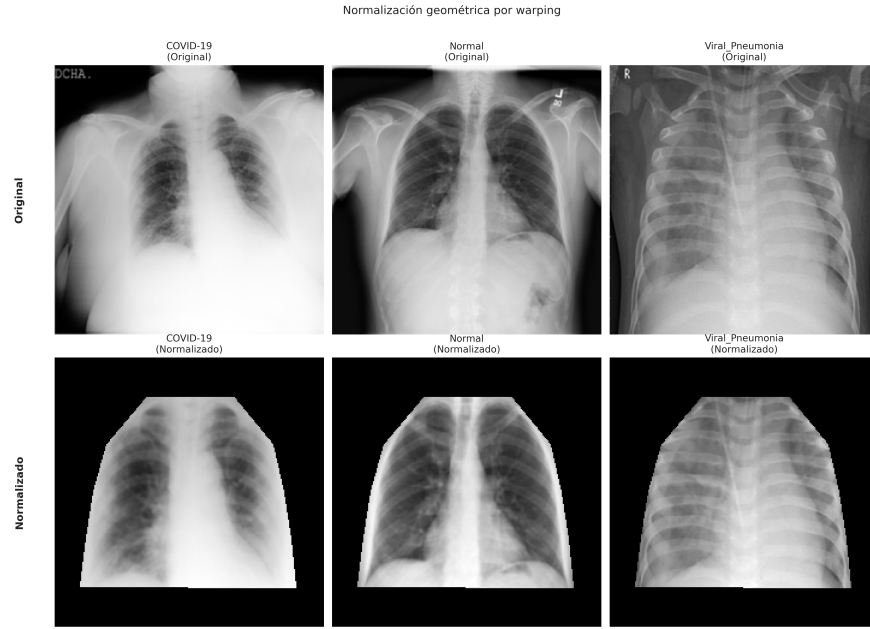
El entrenamiento empleó aumento de datos (volteo horizontal, rotaciones  $\pm 10^\circ$ , transformaciones afines leves), optimizador AdamW [10] con tasa inicial  $1 \times 10^{-4}$ , y parada temprana basada en F1-Macro del conjunto de validación.

## 2.9. Métricas de Evaluación

Para evaluar el sistema se utilizaron las siguientes métricas:

- **Error de landmarks:** Distancia euclidiana promedio entre landmarks predichos y anotados (en píxeles).
- **Accuracy:** Proporción de clasificaciones correctas sobre el total.
- **F1-Macro:** Promedio aritmético del F1-Score de cada clase, dando igual peso a todas independientemente de su frecuencia.
- **Precisión y Sensibilidad:** Por categoría.





**Figura 5.** Ejemplos de normalización geométrica para las tres categorías. Fila superior: radiografías originales de COVID-19, Normal y Neumonía Viral mostrando variaciones de tamaño, posición y orientación. Fila inferior: las mismas imágenes después del proceso de warping, donde la región pulmonar ha sido transformada a la forma estándar, eliminando variaciones no patológicas y enfocando el análisis en la silueta pulmonar normalizada.

### 3. Resultados

Esta sección presenta los resultados experimentales organizados en tres aspectos fundamentales: (1) precisión de localización de landmarks, (2) efecto de la normalización geométrica en la clasificación, y (3) robustez del sistema.

#### 3.1. Precisión de Localización de Landmarks

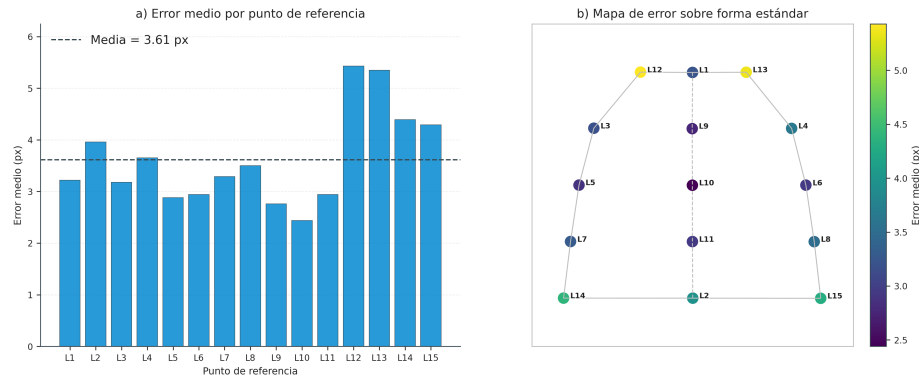
El ensemble de cuatro modelos con Test-Time Augmentation alcanzó un error medio de **3.61 píxeles** en imágenes de  $224 \times 224$  píxeles, lo que representa una mejora del 10.6 % respecto al mejor modelo individual (4.04 píxeles). Este error equivale al 1.6 % del tamaño de la imagen, precisión suficiente para el proceso de normalización geométrica sin introducir distorsiones significativas.

El análisis por landmark reveló un patrón sistemático (Figura 6): mayor precisión en puntos del eje central (2.44–2.94 píxeles) debido a la clara definición de la línea media vertebral, y menor precisión en esquinas superiores (5.35–5.43 píxeles) donde los límites pulmonares son menos nítidos.

**Cuadro 1.** Precisión de detección de landmarks. El error se mide en píxeles sobre imágenes de  $224 \times 224$ .

Configuración	Error promedio	Error mediano	Mejora
Mejor modelo individual	4.04 px	—	—
Ensemble (4 modelos) + TTA	3.61 px	3.07 px	10.6 %

Por categoría, el error fue consistente: Normal (3.22 px), COVID-19 (3.93 px), Neumonía Viral (4.11 px), demostrando robustez del modelo ante diferentes presentaciones patológicas.



**Figura 6.** Análisis del error de predicción por landmark. (a) Gráfica de barras mostrando el error medio en píxeles para cada punto L1–L15, con línea punteada indicando la media global de 3.61 px. Los landmarks L12 y L13 (esquinas superiores) presentan mayor error ( $\sim 5.4$  px), mientras L10 y L11 (eje central) son los más precisos ( $\sim 2.5$  px). (b) Mapa de calor sobre la forma estándar donde el color representa el error medio según la escala de 2.5–5.0 px.

### 3.2. Efecto de la Normalización en la Clasificación

El clasificador entrenado sobre imágenes normalizadas alcanzó los siguientes resultados en el conjunto de prueba de 1,895 imágenes:

- **Accuracy:** 98.10 % (1,859/1,895 correctas)
- **F1-Macro:** 97.17 %
- **F1-Weighted:** 98.09 %

La Tabla 2 presenta el rendimiento por categoría. La clase Normal obtiene el mejor rendimiento (F1: 98.60 %) seguida de COVID-19 (F1: 97.76 %) y Neumonía Viral (F1: 95.15 %), esta última con menor rendimiento debido a su menor representación en el dataset.

**Cuadro 2.** Rendimiento del clasificador por categoría.

Categoría	Precisión	Sensibilidad	F1-Score	Muestras
COVID-19	99.09 %	96.46 %	97.76 %	452
Normal	97.84 %	99.37 %	98.60 %	1,274
Neumonía Viral	97.52 %	92.90 %	95.15 %	169

Para evaluar específicamente el efecto de la normalización geométrica, se realizó una comparación controlada entre tres configuraciones de preprocesamiento, todas utilizando SAHS para mejora de contraste (Tabla 3):

**Cuadro 3.** Comparación de configuraciones de preprocesamiento. Todas utilizan SAHS.

Configuración	Accuracy	F1-Macro	Diferencia
Original + SAHS	98.68 %	97.75 %	—
<b>Warped + SAHS (propuesto)</b>	<b>98.10 %</b>	<b>97.17 %</b>	-0.58 pp
Cropped (12 %) + SAHS	97.68 %	96.66 %	-1.00 pp

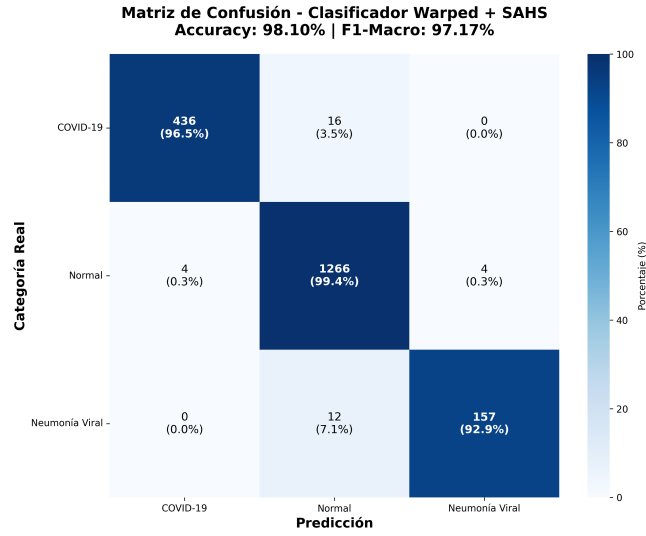
Los resultados revelan hallazgos importantes:

1. **Rendimiento comparable:** El sistema con normalización geométrica (Warped+SAHS) alcanza accuracy apenas 0.58 puntos porcentuales inferior a las imágenes originales.
2. **Posibles características espurias:** Las imágenes originales podrían estar utilizando artefactos hospitalarios en los bordes (etiquetas, marcadores) como “atajos” para la clasificación.
3. **Importancia de la normalización vs. simple recorte:** El recorte de bordes sin normalización (Cropped+SAHS) produce el peor resultado (97.68 %), 1 punto porcentual inferior a las originales, demostrando que la normalización geométrica proporciona un mecanismo efectivo para compensar la pérdida de información periférica.

Esto sugiere que el sistema propuesto aprende características patológicas genuinas de la región pulmonar, no artefactos de los bordes.

### 3.3. Robustez del Sistema

La matriz de confusión (Figura 7) revela patrones de error específicos. La confusión principal ocurre entre Neumonía Viral y Normal (12 casos, 7.1 % de casos Viral), seguida de COVID-19 y Normal (16 casos, 3.5 % de casos COVID). Notablemente, ningún caso de COVID-19 fue confundido con Neumonía Viral ni viceversa, lo cual es clínicamente relevante dado que requieren tratamientos diferentes.



**Figura 7.** Matriz de confusión del clasificador Warped + SAHS (Accuracy: 98.10 %, F1-Macro: 97.17 %). La diagonal muestra las clasificaciones correctas: COVID-19 (436 casos, 96.5 %), Normal (1,266 casos, 99.4 %) y Neumonía Viral (157 casos, 92.9 %). Los errores principales son: 16 casos de COVID-19 clasificados como Normal (3.5 %) y 12 casos de Neumonía Viral como Normal (7.1 %). Notablemente, no hay confusión entre COVID-19 y Neumonía Viral.

La estabilidad del modelo fue evaluada mediante múltiples semillas aleatorias, obteniendo una desviación estándar de  $\pm 0.35$  puntos porcentuales en accuracy, indicando alta reproducibilidad de los resultados.

### 3.4. Comparación con la Literatura

La Tabla 4 compara el sistema propuesto con trabajos relacionados. El sistema alcanza rendimiento competitivo destacando que utiliza un dataset significativamente más grande (15,153 imágenes), clasifica 3 clases (más difícil que clasificación binaria), y emplea una arquitectura más ligera (ResNet-18).

**Cuadro 4.** Comparación con trabajos relacionados en clasificación de COVID-19.

Trabajo	Clases Dataset		Arquitectura	Accuracy
<b>Sistema propuesto</b>	3	15,153	ResNet-18 + Warping	<b>98.10 %</b>
Chowdhury et al. [2]	4	423	VGG19	96.58 %
Rahman et al. [14]	3	3,616	VGG16	93.94 %
Ozturk et al. [12]	2	1,127	DarkCovidNet	98.08 %
Narin et al. [11]	2	341	ResNet-50	98.00 %

## 4. Conclusiones

Este trabajo presentó un sistema integral de clasificación de enfermedades pulmonares basado en normalización geométrica mediante landmarks anatómicos. Los resultados experimentales proporcionan evidencia sobre la efectividad del enfoque propuesto.

### 4.1. Evidencia de Mejora

La comparación controlada entre configuraciones de preprocesamiento demuestra que:

- La normalización geométrica permite obtener rendimiento comparable a las imágenes originales (98.10 % vs 98.68 %), con una diferencia de solo 0.58 puntos porcentuales.
- El simple recorte de bordes sin normalización produce resultados inferiores (97.68 %), lo que sugiere que las imágenes originales podrían beneficiarse de artefactos en los bordes.
- El sistema propuesto, al enfocar el análisis exclusivamente en la región pulmonar normalizada, aprende características patológicas genuinas relacionadas con la enfermedad.

### 4.2. Análisis de Métricas

El sistema completo alcanza métricas competitivas con el estado del arte:

- **Accuracy:** 98.10 % sobre 1,895 imágenes de prueba
- **F1-Macro:** 97.17 %, indicando rendimiento equilibrado entre clases
- **Sensibilidad COVID-19:** 96.46 %, detectando correctamente la mayoría de casos
- **Error de landmarks:** 3.61 píxeles, precisión suficiente para normalización efectiva

La matriz de confusión muestra que los errores principales ocurren entre categorías clínicamente relacionadas (Neumonía Viral/Normal, COVID-19/Normal), mientras que COVID-19 y Neumonía Viral nunca se confunden entre sí.

### 4.3. Contribuciones Principales

Las contribuciones de este trabajo son:

1. Un modelo robusto de predicción de landmarks que alcanza 3.61 píxeles de error mediante ensemble y TTA.
2. Un sistema completo de normalización geométrica integrando GPA, triangulación de Delaunay y warping afín por partes.
3. Evidencia experimental de que la normalización geométrica permite aprender características patológicas genuinas, no artefactos de los bordes.
4. Una metodología reproducible validada sobre un dataset público de 15,153 imágenes.

#### 4.4. Trabajo Futuro

Como direcciones de trabajo futuro se proponen:

- **Validación externa:** Evaluar el sistema en datasets de diferentes instituciones para verificar generalización.
- **Clasificación binaria:** Explorar el rendimiento en la tarea más común de detección COVID-19 vs Normal.
- **Extensión a más patologías:** Incorporar otras enfermedades pulmonares como tuberculosis o cáncer.
- **Integración end-to-end:** Explorar el uso de Spatial Transformer Networks para aprendizaje conjunto de normalización y clasificación.

#### Referencias

1. Ayala-Raggi, S.E., Picazo-Castillo, A.E., Barreto-Flores, A., Portillo-Robledo, J.F.: Synergizing chest x-ray image normalization and discriminative feature selection for efficient and automatic covid-19 recognition. In: Pattern Recognition. ACPR 2023. Lecture Notes in Computer Science, vol. 14407, pp. 224–238. Springer (2023). [https://doi.org/10.1007/978-3-031-47637-2\\_17](https://doi.org/10.1007/978-3-031-47637-2_17)
2. Chowdhury, M.E., Rahman, T., Khandakar, A., Mazhar, R., Kadir, M.A., Mahbub, Z.B., Islam, K.R., Khan, M.S., Iqbal, A., Al Emadi, N., et al.: Can ai help in screening viral and covid-19 pneumonia? IEEE Access **8**, 132665–132676 (2020)
3. Delaunay, B.: Sur la sphère vide. Bulletin de l'Académie des Sciences de l'URSS, Classe des Sciences Mathématiques et Naturelles **6**, 793–800 (1934)
4. Deng, J., Dong, W., Socher, R., Li, L.J., Li, K., Fei-Fei, L.: Imagenet: A large-scale hierarchical image database. In: Proceedings of the IEEE conference on computer vision and pattern recognition. pp. 248–255 (2009)
5. Feng, Z.H., Kittler, J., Awais, M., Huber, P., Wu, X.J.: Wing loss for robust facial landmark localisation with convolutional neural networks. In: Proceedings of the IEEE conference on computer vision and pattern recognition. pp. 2235–2245 (2018)
6. Gower, J.C.: Generalized procrustes analysis. Psychometrika **40**(1), 33–51 (1975)
7. He, K., Zhang, X., Ren, S., Sun, J.: Deep residual learning for image recognition. In: Proceedings of the IEEE conference on computer vision and pattern recognition. pp. 770–778 (2016)
8. Hou, Q., Zhou, D., Feng, J.: Coordinate attention for efficient mobile network design. In: Proceedings of the IEEE/CVF conference on computer vision and pattern recognition. pp. 13713–13722 (2021)
9. Jaderberg, M., Simonyan, K., Zisserman, A., Kavukcuoglu, K.: Spatial transformer networks. In: Advances in neural information processing systems. vol. 28 (2015)
10. Loshchilov, I., Hutter, F.: Decoupled weight decay regularization. In: International Conference on Learning Representations (2019)
11. Narin, A., Kaya, C., Pamuk, Z.: Automatic detection of coronavirus disease (covid-19) using x-ray images and deep convolutional neural networks. Pattern Analysis and Applications **24**(3), 1207–1220 (2021)
12. Öztürk, T., Talo, M., Yildirim, E.A., Baloglu, U.B., Yildirim, O., Acharya, U.R.: Automated detection of covid-19 cases using deep neural networks with x-ray images. Computers in Biology and Medicine **121**, 103792 (2020)

13. Picazo-Castillo, A.E., Ayala-Raggi, S.E., Altamirano-Robles, L., Barreto-Flores, A., Portillo-Robledo, J.F.: Comparative study of lung image representations for automated pneumonia recognition. *International Journal of Combinatorial Optimization Problems and Informatics* **15**(5), 193–201 (2024). <https://doi.org/10.61467/2007.1558.2024.v15i5.578>
14. Rahman, T., Khandakar, A., Qiblawey, Y., Tahir, A., Kiranyaz, S., Kashem, S.B.A., Islam, M.T., Al Maadeed, S., Zughaier, S.M., Khan, M.S., et al.: Exploring the effect of image enhancement techniques on covid-19 detection using chest x-ray images. *Computers in Biology and Medicine* **132**, 104319 (2021)
15. Rajpurkar, P., Irvin, J., Zhu, K., Yang, B., Mehta, H., Duan, T., Ding, D., Bagul, A., Langlotz, C., Shpanskaya, K., et al.: Chexnet: Radiologist-level pneumonia detection on chest x-rays with deep learning. *arXiv preprint arXiv:1711.05225* (2017)
16. Wang, L., Lin, Z.Q., Wong, A.: Covid-net: A tailored deep convolutional neural network design for detection of covid-19 cases from chest x-ray images. *Scientific Reports* **10**(1), 19549 (2020)
17. Wolberg, G.: *Digital image warping*. IEEE Computer Society Press, Los Alamitos, CA (1990)
18. World Health Organization: Use of chest imaging in covid-19: a rapid advice guide. WHO Reference Number: WHO/2019-nCoV/Clinical/Radiology\_imaging/2020.1 (2020)
19. Wu, Y., He, K.: Group normalization. In: *Proceedings of the European conference on computer vision (ECCV)*. pp. 3–19 (2018)
20. Zech, J.R., Badgeley, M.A., Liu, M., Costa, A.B., Titano, J.J., Oermann, E.K.: Variable generalization performance of a deep learning model to detect pneumonia in chest radiographs: a cross-sectional study. *PLoS medicine* **15**(11), e1002683 (2018)