



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

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.

Calendarización de actividades

Actividad	Otoño 2023	Primavera 2024	Otoño 2024	Primavera 2025
Evaluación de programas relacionados	Oct - Ene			
Revisión de trabajos relacionados		Ene - Feb		
Adquisición y preprocesamiento de imágenes		Feb - Mar		
Investigación de algoritmos deformables		Mar - May		
Implementación de la segmentación			May - Oct	
Normalización de imágenes			Oct - Dic	
Alineación de imágenes				Ene - Mar
Extracción y selección de características				Mar - Abr
Desarrollo del clasificador				Abr - May
Validación del clasificador				Jun
Análisis comparativo con otros métodos				Jul

Porcentaje de avance de actividades

Objetivo	% Objetivo	% Avance	% Total
Evaluación de programas relacionados	5 %	100 %	5 %
Revisión de trabajos relacionados	5 %	100 %	5 %
Adquisición y preprocesamiento de imágenes	10 %	100 %	10 %
Investigación de algoritmos deformables	5 %	100 %	5 %
Implementación de la segmentación	15 %	100 %	15 %
Normalización de imágenes	20 %	100 %	20 %
Alineación de imágenes	20 %	100 %	20 %
Extracción y selección de características	5 %	100 %	5 %
Desarrollo del clasificador	5 %	0 %	0 %
Validación del clasificador	5 %	0 %	0 %
Análisis comparativo con otros métodos	5 %	0 %	0 %
Total			85 %

1. Introducción

La radiografía de tórax constituye el estudio de imagenología médica más frecuente a nivel mundial, con aproximadamente 2 mil millones de procedimientos realizados anualmente (Bushberg et al., 2011). Esta modalidad diagnóstica es fundamental para la evaluación de patologías pulmonares, cardiovasculares y torácicas, proporcionando información crítica para la toma de decisiones clínicas. La pandemia de COVID-19 intensificó dramáticamente la demanda de análisis radiológico automatizado, evidenciando las limitaciones de los sistemas existentes (World Health Organization, 2020).

1.1. Planteamiento del Problema

El análisis automatizado de radiografías de tórax enfrenta un desafío fundamental: la **variabilidad geométrica inherente** a las condiciones de adquisición. Las diferencias en posicionamiento del paciente, distancia foco-detector, angulación del haz de rayos X y características anatómicas individuales introducen variaciones significativas en la representación espacial de las estructuras torácicas.

Esta variabilidad impacta negativamente en los sistemas de diagnóstico asistido por computadora (CAD), ya que:

1. Los clasificadores aprenden características dependientes de la geometría específica del conjunto de entrenamiento.
2. La robustez ante artefactos de compresión (común en entornos hospitalarios) se ve comprometida.
3. La generalización a datos de diferentes instituciones es limitada (Zech et al., 2018).

1.2. Pregunta de Investigación

El presente trabajo aborda la siguiente pregunta central:

¿Cómo diseñar un sistema de normalización geométrica basado en landmarks anatómicos que mejore la robustez y precisión de clasificadores de patologías pulmonares en radiografías de tórax?

1.3. Enfoque Propuesto

Se propone un pipeline de tres etapas:

1. **Predicción de landmarks:** Localización automática de 15 puntos anatómicos mediante redes neuronales convolucionales (He et al., 2016) con función de pérdida Wing Loss (Feng et al., 2018) y restricciones geométricas.

2. **Normalización geométrica:** Transformación de la imagen a una forma canónica mediante análisis de Procrustes (Gower, 1975) y warping afín por partes.
3. **Clasificación:** Detección de patologías (COVID-19, neumonía viral, normal) sobre imágenes normalizadas.

Este enfoque permite separar la variabilidad geométrica extrínseca de las características patológicas intrínsecas, mejorando tanto la precisión como la robustez del sistema.

En la siguiente sección se formulan las hipótesis cuantificables que guían el desarrollo experimental (Sección 2).

2. Hipótesis

2.1. Hipótesis Principal

Como se establece en la Sección 1, la variabilidad geométrica representa un desafío fundamental para los sistemas CAD. Se plantea que la incorporación de restricciones geométricas anatómicas en la función de pérdida de redes neuronales convolucionales (LeCun et al., 1998), combinada con normalización espacial mediante warping afín por partes, mejorará significativamente la robustez de clasificadores de patologías pulmonares ante artefactos de compresión y variabilidad geométrica.

2.2. Variables de Investigación

Variables Independientes:

- Función de pérdida: MSE, Wing Loss (Feng et al., 2018), Wing Loss + restricciones geométricas
- Estrategia de entrenamiento: backbone congelado vs. fine-tuning completo
- Porcentaje de cobertura (fill rate) del warping: 47 %, 96 %, 99 % (ver definición en Ecuación 7)

Variables Dependientes:

- Error euclidiano de predicción de landmarks (píxeles)
- Accuracy de clasificación (%)
- Degradación de accuracy bajo compresión JPEG (%)
- Gap de generalización en validación cruzada (%)

2.3. Predicciones Cuantificables

1. **Predicción de landmarks:** El ensemble de modelos alcanzará un error medio menor a 5 píxeles en imágenes de 224×224 .
2. **Robustez a compresión:** El clasificador entrenado con imágenes normalizadas presentará al menos $5 \times$ menor degradación bajo compresión JPEG Q50 comparado con el baseline sin normalización.
3. **Generalización:** El gap de accuracy en validación cruzada (modelo evaluado en dataset diferente al de entrenamiento) será al menos $2 \times$ menor para el modelo con normalización geométrica.

2.4. Métrica de Evaluación Principal

El error de localización de landmarks se define como:

$$E_k = \sqrt{(\hat{x}_k - x_k)^2 + (\hat{y}_k - y_k)^2} \quad (1)$$

donde (\hat{x}_k, \hat{y}_k) son las coordenadas predichas y (x_k, y_k) las coordenadas de referencia (ground truth) para el landmark k .

La relevancia de estas predicciones se fundamenta en necesidades clínicas concretas, como se detalla en la Sección 3.

3. Justificación

3.1. Relevancia Clínica

La pandemia de COVID-19 evidenció dramáticamente la brecha entre la demanda de análisis radiológico y la capacidad disponible de especialistas (World Health Organization, 2020). En el pico de la crisis sanitaria, los sistemas de salud enfrentaron:

- Incremento exponencial en volumen de radiografías de tórax
- Necesidad de triaje rápido para priorización de pacientes
- Variabilidad en equipos y protocolos entre instituciones
- Compresión de imágenes para transmisión y almacenamiento

Los sistemas de diagnóstico asistido por computadora (CAD), como COVID-Net (Wang et al., 2020) y CheXNet (Rajpurkar et al., 2017), ofrecen una solución potencial, pero su efectividad depende críticamente de la robustez ante las condiciones reales de operación hospitalaria.

3.2. Contribuciones Científicas

Este trabajo aporta las siguientes contribuciones:

1. **Robustez mejorada a artefactos:** La normalización geométrica proporciona $5.3\times$ mejor robustez a compresión JPEG Q50 (3.06 % vs 16.14 % degradación), alcanzando hasta $30\times$ con warping de 47 % fill rate, crítico para deployment en entornos hospitalarios donde la compresión es inevitable.
2. **Función de pérdida geométrica multi-componente:** Integración de Wing Loss (Feng et al., 2018) con restricciones de alineación central y simetría bilateral, logrando error de localización de 3.71 píxeles en 15 landmarks anatómicos.
3. **Análisis del mecanismo de robustez:** Experimento de control que identifica dos componentes causales de la mejora: aproximadamente 75 % atribuible a reducción de información (regularización implícita) y 25 % adicional a normalización geométrica propiamente dicha.
4. **Validación rigurosa con limitaciones documentadas:** Evaluación en dataset externo FedCOVIDx (8,482 muestras de múltiples instituciones) con documentación honesta del domain shift (Zech et al., 2018). La accuracy de 53–55 % en datos externos (comparable al modelo original sin normalización) confirma que el domain shift cross-institucional persiste, consistente con la literatura en imágenes médicas.
5. **Pipeline reproducible:** Implementación completa con código fuente, configuraciones y checkpoints disponibles para replicación.

3.3. Impacto Potencial

La normalización geométrica propuesta puede servir como:

- Preprocesamiento estándar para sistemas CAD pulmonares
- Base para segmentación automática mediante Active Shape Models
- Método de regularización implícita para mejorar generalización
- Técnica de reducción de sensibilidad a artefactos de compresión

3.4. Alcance y Limitaciones

Es importante delimitar que este trabajo:

- Se enfoca en radiografías PA de tórax (proyección posteroanterior)
- Valida dentro del dominio de entrenamiento con alta precisión (99.10 % accuracy)
- Documenta honestamente las limitaciones de generalización cross-institucional
- No pretende reemplazar el diagnóstico médico profesional

Para implementar la solución propuesta, es necesario establecer los fundamentos teóricos que se presentan en la Sección 4.

4. Marco Teórico

4.1. Radiografías de Tórax

La radiografía de tórax se fundamenta en la atenuación diferencial de rayos X al atravesar tejidos de distinta densidad (Bushberg et al., 2011). La intensidad transmitida $I(x)$ sigue la ley de Beer-Lambert:

$$I(x) = I_0 \exp \left(- \int_0^x \mu(s) ds \right) \quad (2)$$

donde I_0 es la intensidad incidente, $\mu(s)$ el coeficiente de atenuación lineal (dependiente del tejido), y x la distancia recorrida.

Los coeficientes típicos de atenuación son:

- Aire alveolar: $\mu \approx 0,0001 \text{ cm}^{-1}$
- Tejidos blandos: $\mu \approx 0,20 \text{ cm}^{-1}$
- Hueso cortical: $\mu \approx 0,50 \text{ cm}^{-1}$

4.1.1. Landmarks Anatómicos

Se definen 15 puntos de referencia que caracterizan la geometría torácica:

- **Eje central** (L1, L2): Define la línea media vertical
- **Puntos centrales** (L9, L10, L11): Dividen el eje en cuartos
- **Pares bilaterales**: (L3-L4), (L5-L6), (L7-L8), (L12-L13), (L14-L15)

Estos landmarks presentan propiedades geométricas verificables: los puntos centrales se ubican exactamente en $t \in \{0,25, 0,50, 0,75\}$ a lo largo del eje, con desviación menor a 1.5 píxeles.

4.2. Redes Neuronales Convolucionales

Las CNNs (LeCun et al., 1998) realizan extracción jerárquica de características mediante la operación de convolución discreta:

$$Y[i, j] = \sum_{m=0}^{M-1} \sum_{n=0}^{N-1} X[i + m, j + n] \cdot W[m, n] + b \quad (3)$$

donde X es la entrada, W el kernel de convolución, y b el sesgo.

4.2.1. Arquitectura ResNet

Las redes residuales (He et al., 2016) introducen conexiones de salto (skip connections) que permiten entrenar redes más profundas:

$$\mathbf{y} = \mathcal{F}(\mathbf{x}, \{W_i\}) + \mathbf{x} \quad (4)$$

donde \mathcal{F} representa las capas residuales y \mathbf{x} la conexión directa.

ResNet-18, con 11.7 millones de parámetros, ofrece un balance óptimo entre capacidad representacional y eficiencia computacional para tareas de regresión de coordenadas.

4.2.2. Transfer Learning

El preentrenamiento en ImageNet (1.2 millones de imágenes, 1000 clases) proporciona:

- Filtros de bajo nivel (bordes, texturas) altamente transferibles
- Inicialización superior a pesos aleatorios
- Convergencia más rápida con menos datos médicos

Para tareas con menos de 10,000 imágenes etiquetadas, el transfer learning mejora consistentemente el rendimiento comparado con entrenamiento desde cero (Yosinski et al., 2014).

4.3. Funciones de Pérdida para Regresión de Landmarks

La regresión de coordenadas de landmarks presenta desafíos específicos que requieren funciones de pérdida especializadas.

4.3.1. Wing Loss

Para la localización precisa de landmarks, la función Wing Loss (Feng et al., 2018) ofrece mejor convergencia que el error cuadrático medio (MSE) al balancear la sensibilidad a errores pequeños y grandes:

$$\mathcal{L}_{\text{wing}}(y, \hat{y}) = \begin{cases} w \ln \left(1 + \frac{|y - \hat{y}|}{\epsilon} \right) & \text{si } |y - \hat{y}| < w \\ |y - \hat{y}| - C & \text{en otro caso} \end{cases} \quad (5)$$

donde w controla el rango no lineal (típicamente $w = 10$), ϵ previene división por cero ($\epsilon = 2$), y $C = w - w \ln(1 + w/\epsilon)$ asegura continuidad en $|y - \hat{y}| = w$.

Esta formulación penaliza más fuertemente los errores pequeños (región logarítmica) mientras mantiene comportamiento lineal para outliers, resultando en predicciones más precisas cerca del ground truth.

4.4. Normalización Geométrica

4.4.1. Análisis de Procrustes Generalizado

La alineación de formas utiliza el Análisis de Procrustes Generalizado (GPA) (Gower, 1975; Dryden and Mardia, 2016) para encontrar la transformación óptima entre conjuntos de landmarks. Dadas dos configuraciones de puntos X e Y , la rotación óptima se obtiene mediante:

$$R^* = VU^T \quad \text{donde} \quad \text{SVD}(X^T Y) = U\Sigma V^T \quad (6)$$

La transformación afín completa incluye traslación, escala y rotación, minimizando la distancia de Procrustes entre la configuración observada y una forma de referencia canónica.

4.4.2. Warping Afín por Partes

El warping afín por partes divide la imagen en regiones triangulares (triangulación de Delaunay) y aplica transformaciones afines locales $T_i(\mathbf{x}) = A_i\mathbf{x} + \mathbf{b}_i$ a cada triángulo. Cada matriz $A_i \in \mathbb{R}^{2 \times 2}$ se determina por las correspondencias de los tres vértices del triángulo entre la configuración original y la normalizada.

El *fill rate* (porcentaje de cobertura) cuantifica la fracción del área de salida con información válida después del warping:

$$\text{fill_rate} = 1 - \frac{\text{píxeles_negros}}{\text{total_píxeles}} \quad (7)$$

Valores típicos son 47 % (warping conservativo, solo región central) y 96 % (cobertura extendida con interpolación en bordes).

Referencias

- Bushberg, J. T., Seibert, J. A., Leidholdt Jr, E. M., and Boone, J. M. (2011). *The essential physics of medical imaging*. Lippincott Williams & Wilkins.
- Dryden, I. L. and Mardia, K. V. (2016). *Statistical shape analysis: with applications in R*. John Wiley & Sons.
- Feng, Z.-H., Kittler, J., Awais, M., Huber, P., and Wu, X.-J. (2018). Wing loss for robust facial landmark localisation with convolutional neural networks. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 2235–2245.
- Gower, J. C. (1975). Generalized procrustes analysis. *Psychometrika*, 40(1):33–51.
- He, K., Zhang, X., Ren, S., and Sun, J. (2016). Deep residual learning for image recognition. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 770–778.
- LeCun, Y., Bottou, L., Bengio, Y., and Haffner, P. (1998). Gradient-based learning applied to document recognition. *Proceedings of the IEEE*, 86(11):2278–2324.
- Rajpurkar, P., Irvin, J., Zhu, K., Yang, B., Mehta, H., Duan, T., Ding, D., Bagul, A., Langlotz, C., Shpanskaya, K., et al. (2017). Chexnet: Radiologist-level pneumonia detection on chest x-rays with deep learning. *arXiv preprint arXiv:1711.05225*.
- Wang, L., Lin, Z. Q., and Wong, A. (2020). 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.
- World Health Organization (2020). Use of chest imaging in covid-19: a rapid advice guide. *WHO Reference Number: WHO/2019-nCoV/Clinical/Radiology_imaging/2020.1*.
- Yosinski, J., Clune, J., Bengio, Y., and Lipson, H. (2014). How transferable are features in deep neural networks? In *Advances in neural information processing systems*, volume 27.
- Zech, J. R., Badgeley, M. A., Liu, M., Costa, A. B., Titano, J. J., and Oermann, E. K. (2018). Variable generalization performance of a deep learning model to detect pneumonia in chest radiographs: a cross-sectional study. *PLoS medicine*, 15(11):e1002683.