

Clasificación de pacientes con neumonía a partir de rayos X de tórax

Alejandra Barrera
Universidad de los Andes
Bogotá D.C.

ma.barrera12@uniandes.edu.co

Jessica Castillo
Universidad de los Andes
Bogotá D.C.

j1.castillo@uniandes.edu.co

Juan David García
Universidad de los Andes
Bogotá D.C.

jd.garcia20@uniandes.edu.co

1.. Introducción

La neumonía es una infección del parénquima pulmonar causado por la bacteria *Streptococcus pneumoniae*. Esta infección inflama los alveolos en uno o dos pulmones, los cuales se llenan de fluido causando dificultad respiratoria y llegando a provocar asfixia y muerte. Para el año 2015, 920,136 niños menores a 5 años murieron debido a esta patología en todo el mundo siendo la principal causa de muerte infecciosa y cobrando más vidas que el SIDA, la malaria y el sarampión combinados [4]. De acuerdo con la Organización Mundial de la Salud, la neumonía es la 8va causa de muerte en los Estados Unidos y la cuarta causa de muerte en pacientes mayores de 65 años. Cuando se contrae neumonía, la infección comienza con la aparición de fiebre y síntomas respiratorios los cuales van acompañados de infiltrados pulmonares que se pueden visualizar en una radiografía de tórax [8].

Actualmente, la neumonía es diagnosticada mediante un análisis de la historia clínica del paciente en lo que refiere a enfermedades del sistema respiratorio, signos vitales, exámenes de laboratorio y una revisión de la radiografía de tórax realizado por un especialista. Generalmente, cuando hay presencia de neumonía, la zona del pulmón que se ve afectada, se ve más opaca que el resto del tejido [8]. Sin embargo, el diagnóstico a partir de la radiografía de tórax resulta ser complicado debido a que se pueden presentar distintas condiciones en los pulmones tales como edemas pulmonares, sangrado, actelestasia entre otros [8]. En consecuencia, el diagnóstico temprano y automatizado mediante el análisis de imágenes médicas se considera un reto debido a la variabilidad de las radiografías y criterios obtenidos. Generalmente, las radiografías pueden ser engañosas y requieren de un especialista con experiencia. Además, en

enfermedades como la neumonía el diagnóstico es basado en la opinión del radiólogo y de cómo este interprete una radiografía [4].

Debido a lo anterior, se plantea el desarrollo de un algoritmo de clasificación de neumonía basado en el análisis de rayos x de tórax, ya que esto permitiría una detección médica temprana y, en consecuencia, un mejor tratamiento para la población afectada. Por esto, al implementar dicho algoritmo sin sesgos de experiencia o percepción se espera que sea una herramienta que ayude al profesional a analizar, enfocar y resaltar ciertos detalles que el ojo humano pudo haber pasado por alto. Un aspecto importante es que los algoritmos pueden aprender de un gran número de imágenes en bases de datos que constantemente se siguen actualizando. Por esto, se puede reducir el error humano. Actualmente, existen varios *challenges* e investigaciones por parte de universidades o agencias de salud que incentivan la implementación de la inteligencia artificial (AI) a este tipo de problemas biomédicos [2].

Por tanto, el fundamento del problema se basa en la construcción de un algoritmo que logre clasificar las imágenes de radiografías, particularmente radiografías de tórax, en aquellas que posean o no neumonía. Específicamente, el objetivo es que dicho algoritmo automáticamente localice las regiones de los pulmones donde hay opacidad, siendo este el principal signo de neumonía. Por tanto, el problema de análisis y procesamiento de imágenes médicas es de clasificación de forma que se determine si un paciente tiene o no signos de la patología. Así, el objetivo es que los resultados se vean reflejados en *bounding boxes* para cada imagen de prueba, y un nivel de confianza de dicho cuadro. Ahora bien, el problema está diseñado para que se entrene un modelo utilizando la información de los bounding boxes (BB)

de imágenes con neumonía, los cuales son provistos, y los BB encontrados experimentalmente; y se aplique este modelo a imágenes de prueba con el fin de evaluar la eficacia del algoritmo.

2.. Trabajo relacionado

De acuerdo con la literatura, se encontró que en distintos estudios se parte de una base de datos anotada manualmente en donde los bounding boxes (cuadros delimitadores) corresponden a nódulos pulmonares derivados producidos por la patología. Uno de los métodos más ampliamente usados para este problema es aquel en donde se implementan redes neuronales convolucionales basadas en regiones (R-CNN), por sus siglas en inglés. Este método se basa en un proceso de múltiples etapas en donde se busca extraer características de cada una de las regiones a clasificar las cuales han sido anotadas previamente. La siguiente etapa de este algoritmo consiste en entrenar simultáneamente un modelo de regresión lineal para corregir el bounding-box que fue predicha a partir de compensaciones aprendidas por el algoritmo. Sin embargo, en la literatura se explica que este método es muy costoso computacionalmente debido al enfoque en múltiples etapas sin un cómputo compartido [6]. Actualmente, el uso de R-CNN o alguna variación de estas es el método que se utiliza hoy en día. En los artículos relacionados con redes neuronales, generalmente, los autores no realizan algún tipo de pre-procesamiento de las imágenes, por ejemplo, en términos de filtrado. Por último, los artículos exponen las metodologías, sin embargo, no hay código explícito. No obstante, en la página oficial del reto ('RSNA Pneumonia Detection Challenge') se encuentran múltiples soluciones al problema, o a parte de él, denominados "kernels". Estos son cuadernos de Jupyter que sí contienen código, y en su mayoría, las soluciones propuestas también están relacionadas con redes neuronales [1].

Por otro lado, una de las soluciones que se encontraron para este problema, en donde no se hacía uso de R-CNN, es la propuesta de Saad *et al* que consiste en el uso de un filtrado denominado *canny edge* para mejorar la segmentación del pulmón. En dicho trabajo se desarrolló el problema a partir de una segmentación de la región de interés, en donde partiendo de la detección de los bordes del pulmón se segmentaron mediante operaciones morfológicas. En primer lugar, los autores superan el problema de la pobre segmentación que se había observado utilizando algoritmos de segmentación de bordes como el de Sobel, Prewitt y Laplace, ya que estos son filtros pasa-altas y sensibles al ruido de la imagen. Teniendo esto en cuenta, propusieron el uso de un algoritmo de detección *canny edge* y, realizaron operaciones morfológicas de dilatación y erosión para obtener un mejor resultado[9].

Por último, se encontró un trabajo realizado por Sharma *et al* en donde los investigadores basaron el desarrollo de su algoritmo en la Umbralización de Otsu. Para esto, en la etapa de preprocesamiento, los desarrolladores comenzaron por llevar a cabo un redimensionamiento de las imágenes en donde todas quedaron de un tamaño de 800 x 800. Seguido a esto se llevó a cabo una equalización del histograma, para luego recortar la región inferior del abdomen. Como últimos pasos llevaron a cabo una delimitación de la región que comprende a los pulmones, para implementar la umbralización de Otsu sobre esta imagen resultante y obtener las regiones del pulmón que se ven afectadas por la neumonía [12].

3.. Propuesta

Partiendo del problema planteado, es decir, la clasificación de imágenes de radiografía entre positivas (que contienen neumonía) o negativas (que no poseen neumonía) se propuso desarrollar el algoritmo a partir de dos acercamientos como se explica a continuación:

3.1.. Clasificación a partir de bounding boxes

Un primer acercamiento consistió en clasificar las imágenes como positivo para neumonía o no, a partir de *bounding boxes*. Para esto, se procedió a obtener los *bounding boxes* de las imágenes que son positivas para neumonía teniendo en cuenta las coordenadas X, Y de la esquina superior izquierda como también el *ancho* y el *alto* del bounding box. Esta información se obtuvo a partir del archivo de anotaciones de la base de datos. Seguido a esto, se generaron los *bounding boxes* para las imágenes que no estaban anotadas como positivas para neumonía. Para esto, se encontró experimentalmente que, para la mayoría de imágenes, el bounding box se encontraba en la zona del pulmón izquierdo con coordenadas $X = 300, Y = 300, ancho = 170$ y $alto = 450$. Como último paso en esta etapa se procedió a redimensionar los bounding boxes a un tamaño de 500×500 con el fin de obtener imágenes de tamaño homogéneo.

Ahora bien, las imágenes de pacientes que tienen neumonía presentan opacidades en las regiones de los pulmones afectadas por la patología como se muestra en la Figura 1a. En cambio, para las imágenes de pacientes que no presentan neumonía los pulmones se ven radiolúcidos como se muestra en la Figura 1b. Por lo tanto, a la hora de generar el bounding box sobre los pulmones de pacientes sanos se evidencia una diferencia marcada a nivel de intensidades de la imagen comparados con los bounding boxes de los pacientes con neumonía como se muestra en la Figura 3. Debido a esto, se encontró que el mejor descriptor para resolver el

problema de clasificación es el histograma de intensidades de los *bounding boxes* redimensionados.

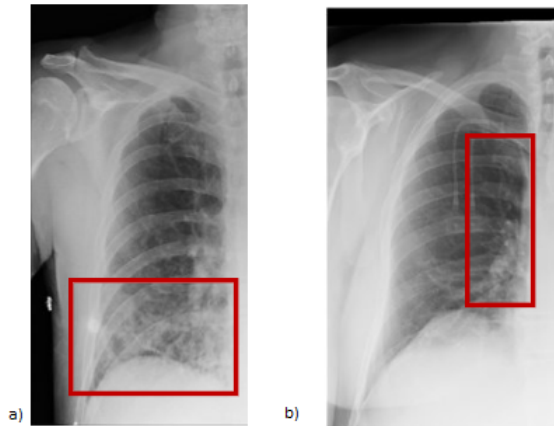


Figura 1: a) Radiografía de un paciente que presenta neumonía. b) Radiografía de un paciente sano.

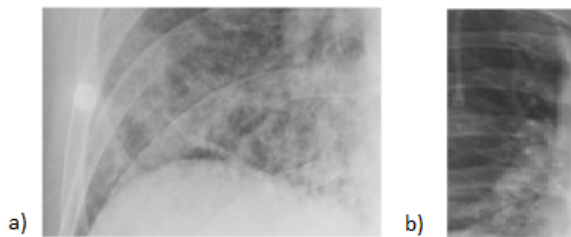


Figura 2: a) Bounding box de un paciente que presenta neumonía. b) Bounding box de un paciente sano.

Por último, una vez obtenidos los histogramas de intensidades de los bounding boxes tanto para las imágenes de pacientes enfermos como para los pacientes sanos, se clasificaron las imágenes usando bosques aleatorios. El procedimiento para el método de clasificación a partir de *bounding boxes* se muestra a continuación:

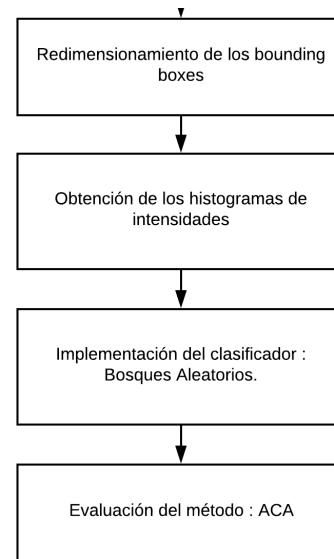
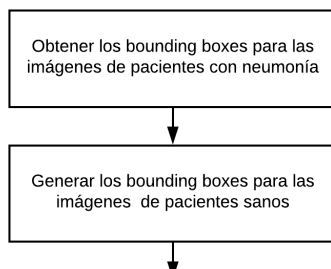
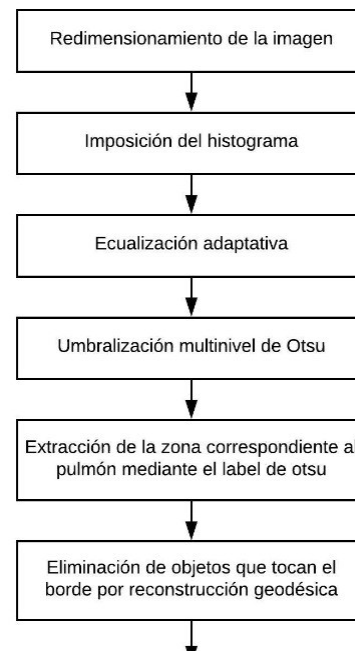


Figura 3: Diagrama de bloques para el método de clasificación a partir de *bounding boxes*.

3.2.. Clasificación a partir de la imagen de radiografía completa

Por otro lado, se realizó otra metodología de clasificación en donde se realiza la extracción del pulmón y, se aplica el mismo procedimiento de la sección anterior para obtener el mejor ACA como se muestra en el siguiente diagrama de bloques.



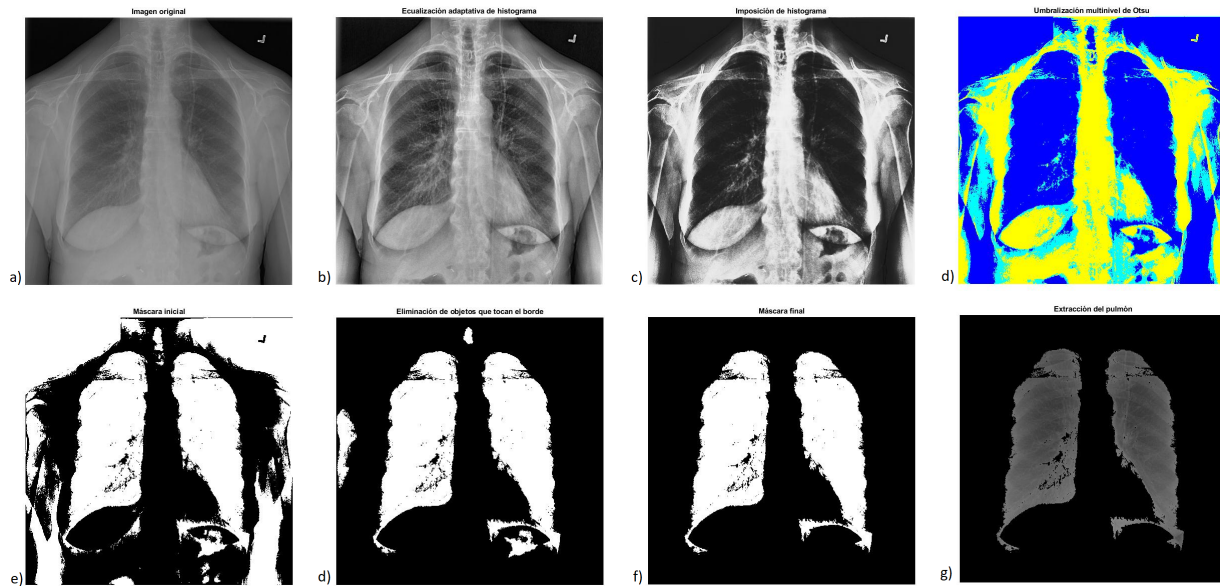


Figura 4: *a) Imagen original, b) Ecualización adaptativa, c) Imposición del histograma, d) Umbralización multinivel de Otsu, e) Máscara inicial obtenida a partir de los labels de la umbralización, f) Eliminación de los objetos que tocan el borde de la imagen, g) Máscara final, h) Extracción del pulmón*

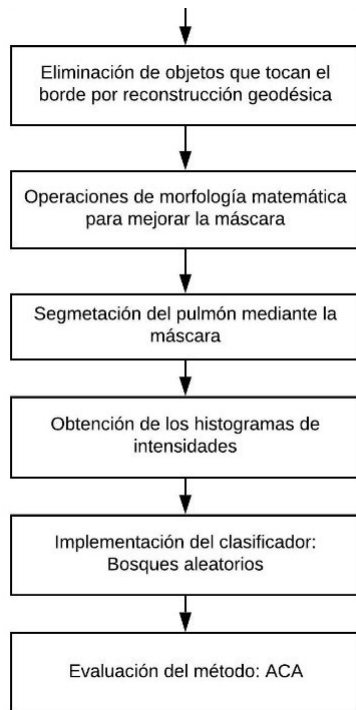


Figura 5: *Diagrama de bloques para el método de clasificación a partir imagen completa de radiografía.*

- En primer lugar, se redimensionaron las imágenes de 1024×1024 a un tamaño de 800×800 para reducir el gasto computacional del algoritmo. Posteriormente, se

impuso el histograma de una imagen de alto contraste para equalizar las imágenes de la base de datos. Esto se hizo con el objetivo de lograr una mejor segmentación, ya que algunas imágenes de la base de datos presentan problemas frente a la segmentación del pulmón utilizando la umbralización de Otsu. Esto permitió una mejor diferenciación de los pulmones y los tejidos más radioclaros como huesos y órganos como se muestra en la Figura 4b.

- A partir de la umbralización de *Otsu*, fue posible diferenciar el fondo de la imagen, tejidos aledaños junto con los huesos y los pulmones. Así se identificó, gracias al diagrama de etiquetas obtenido de la umbralización multinivel de *Otsu*, aquellas que correspondían a los pulmones y se asignaron a una nueva matriz para que, en esta nueva imagen, se llevase a cabo la eliminación de aquellos objetos que tocan el borde de la imagen. No obstante, se realizaron operaciones morfológicas para obtener la mejor máscara para la extracción del pulmón. Entre estas operaciones se encuentran una apertura, después una dilatación y su consecuente reconstrucción utilizando la máscara previamente generada en el orden respectivo.
- Una vez obtenidas las regiones del pulmón deseadas, se procedió a obtener los histogramas de intensidades de cada una de estas imágenes resultantes y se implementaron bosques aleatorios para la clasificación de las imágenes.

4.. Experimentos

4.1.. Base de datos

Una vez presentado el problema y definida la motivación y la relevancia de este, se buscó una base de datos a partir de la cual se pudiera implementar un código para la posible solución. Debido a que se trabajó en un problema de clasificación, el método de evaluación del algoritmo creado fue en base a la precisión de clasificación promedio (ACA) y su matriz de confusión teniendo los datos de las anotaciones provistos por el *challenge*.

La base de datos con la que se trabajó se obtuvo de la plataforma Kaggle del 'RSNA Pneumonia Detection Challenge' en la cual se encuentran los siguientes datos e información:

- Dos carpetas 'train' y 'test' en donde se encuentran las correspondientes imágenes en archivos tipo DICOM. Con estas imágenes se llevó a cabo el entrenamiento y la prueba de los algoritmos a desarrollar para resolver el problema de detección [7].
- Para cada imagen, se especifican los 'data fields' en donde se tiene: el 'PatientId' que representa una imagen única, 'X' como la coordenada superior izquierda del *bounding box*, 'Y' como la coordenada superior izquierda del *bounding box*, 'Width' y 'Height' como el ancho y el alto del *bounding box*. Por último el 'Target', de tipo binario, que indica si la imagen muestra evidencia neumonía o no [7].
- En cuanto al tamaño de las imágenes, estas tienen un tamaño de 1024×1024 .

La Figura 6 muestra una radiografía extraída de la base de datos, de un paciente sano y una radiografía de un paciente con neumonía la cual cuenta con los *bounding boxes* para implementar el método de detección.

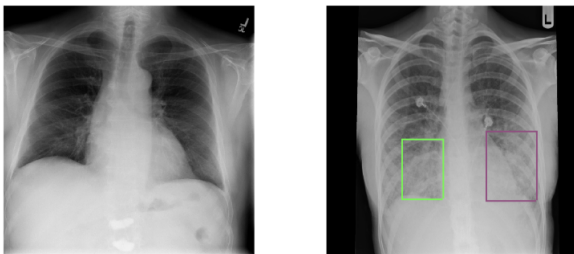


Figura 6: Radiografía de un paciente sano y un paciente con neumonía con sus correspondientes *bounding boxes*.

- Por otro lado, en lo que refiere al número de imágenes, la base de datos cuenta con 26.684 imágenes anotadas,

en donde 6.012 de ellas cuentan con anotaciones positivas de neumonía. El otro grupo de las 20.672 son imágenes de radiografías que no muestran signos de neumonía [6]. Ya que la base de datos se encontraba desordenada y desbalanceada, se organizó en grupos de entrenamiento y prueba. El grupo de entrenamiento constaba de 10.020 imágenes donde 5010 imágenes eran positivas para neumonía y 5010 eran imágenes de pacientes sanos. Para el grupo de prueba se mantuvo la proporción inicial de la base de datos original en donde el 77.5 % de la base de datos eran imágenes de pacientes sanos y el 22.5 % corresponde a imágenes de pacientes con neumonía. Como resultado se obtuvo una carpeta con 4.453 imágenes. Esto es debido a que en situaciones reales médicas no hay un balance de clases.

4.2.. Experimentos de validación

Los experimentos de validación se hicieron sobre el problema de clasificación de *bounding boxes*. Una vez encontrado el mejor descriptor y el mejor clasificador con los mejores parámetros, se utilizaron estos resultados para resolver el problema de clasificación con la imagen completa de radiografía de tórax.

- En primera instancia, se enfocó el interés en encontrar el mejor descriptor mediante el cual se clasifican las imágenes. Para esta etapa se utilizaron histograma de intensidades, histograma de gradientes orientados y textones. Se encontró que mediante la implementación del clasificador *KNN* (*K-Nearest Neighbors*) el mejor ACA utilizando textones fue de 45,31 %, variando el valor de *k-means* que computa los textones. Por otro lado, el mejor ACA obtenido utilizando histogramas de gradientes orientados fue de 38,62 %, variando la cantidad de orientaciones y el tamaño de las celdas. Por último, el mejor ACA utilizando los histogramas de intensidades fue de 80,78 %, en donde se varió el número de bins. Por tanto, ya que el mejor resultado se obtuvo del histograma de intensidades, este descriptor se fijó para la siguiente etapa, así como el mejor valor para el parámetro de bins, es decir, 16.
- El siguiente experimento consistió en encontrar el mejor clasificador. Para ello, se probaron los clasificadores *SVM*, *KNN* y *bosques aleatorios*. Por un lado, la clasificación por *KNN* arrojó que el mejor ACA obtenido fue de 80,78 %, al variar el número de vecinos y encontrar que el mejor fue de 5. En segundo lugar, para el clasificador *SVM* se variaron los kernels, siendo el *RBF* el que arrojó un mejor resultado (ACA = 83,70 %). Por último, los bosques aleatorios mostraron un mejor comportamiento, pues el máximo ACA

alcanzado fue de 85,14 %. Debido a que este clasificador mostró el mejor comportamiento, la variación detallada de sus parámetros se muestra en la Tabla 1.

Tabla 1: Variación de parámetros para los bosques aleatorios. (A) Resultados del ACA en función del número de árboles, dejando constante $B = 0.5$ (fracción de datos para cada árbol) y $C = nbins*3/2$ (aleatoriedad del nodo); (B) resultados del ACA variando la fracción de datos de cada árbol, fijando $A = 200$ y $C = nbins*3/2$; (C) resultados del ACA en función de la aleatoriedad del nodo, fijando $A = 200$ y $B = 0.2$.

(A) No. de árboles		(B) Fracción de datos		(C) Aleatoriedad del nodo	
$B = 0.5$ $C = nbins*3/2$		$A = 200$ $C = nbins*3/2$		$A = 200$ $B = 0.2$	
Valor	ACA (%)	Valor	ACA %	Valor	ACA %
10	82.98	0.2	85.02	nbins*3/1	84.79
30	84.11	0.3	85.01	nbins*3/2	85.02
50	84.61	0.5	84.85	nbins*3/3	85.14
100	84.79	0.7	83.12	nbins*3/8	84.95
200	84.85	0.8	83.76	nbins*3/16	84.72

- Ahora bien, para la clasificación de neumonía tomando en cuenta la imagen de la radiografía completa, se utilizó el mejor descriptor y los mejores parámetros del clasificador seleccionado, es decir, del bosque aleatorio. No obstante, para este proceso se hicieron experimentos relacionados con la extracción del pulmón como umbralización por número de Euler [10], fuzzy c-means [11], operaciones morfológicas, transformación de histogramas, entre otros. Sin embargo, la combinación de métodos que arrojó la mejor segmentación del pulmón fue la expuesta en el diagrama de la Figura 5.

4.3.. Experimentos de prueba

Por un lado, la clasificación de los bounding boxes, usando los mejores parámetros, arrojó la matriz de confusión que se muestra en la Tabla 2, de la que se obtuvo un ACA del 85.14 %.

Tabla 2: Matriz de confusión del método de clasificación de bounding boxes.

Predicción	Anotación		
	Clase	Sano	Neumonía
	Sano	82.85 %	12.57 %
	Neumonía	17.15 %	87.43 %

Por otro lado, la clasificación de las imágenes de las radiografías completas arrojó la matriz de confusión normalizada que se muestra en la tabla 3. De esta se obtuvo un ACA del 70.42 %.

Tabla 3: Matriz de confusión del método de clasificación de la imagen completa.

Predicción	Anotación		
	Clase	Sano	Neumonía
	Sano	69.08 %	28.24 %
	Neumonía	30.92 %	71.76 %

Comparando con otros métodos de clasificación de neumonía, se encontró que en dichas investigaciones se utilizaban tomografías computarizadas (CT) en donde el descriptor para la clasificación son los textones. De ello obtuvieron un ACA del 81 % usando SVM como clasificador [5].

5.. Discusión

Teniendo en cuenta los resultados de los experimentos presentados, se encontró que el mejor descriptor fue el histograma de intensidades. Esto se debe a que la diferencia principal entre una radiografía de un paciente con neumonía y la de un paciente sano son las opacidades que se presentan en las zonas del pulmón afectadas por la patología, como se mostró en la Figura 1. Por otro lado se evidenció que haciendo uso de descriptores de forma y textura la efectividad del método para clasificar las imágenes es bastante baja, ya que a la hora de llevar las imágenes a estos espacios de representación, no existe una diferencia significativa entre las imágenes de pacientes sanos y las imágenes de pacientes con neumonía.

Así, una de las limitaciones principales del método desarrollado es el tipo de imagen diagnóstica. Por ejemplo, una radiografía provee información bidimensional (2D) mientras que una tomografía computarizada (CT, por sus siglas en inglés) provee información tridimensional (3D). Por lo tanto, la información que se puede extraer se ve limitada y, en consecuencia, la precisión del algoritmo. De hecho, se encontró en la literatura que la gran parte de los estudios de clasificación de neumonía se realizan a partir de información de CT. Adicionalmente, de estas imágenes se puede obtener información de la textura de los pulmones y clasificarlas por textones. Por lo que se ve una ventaja significativa de las tomografías frente a las radiografías.

Por otro lado, el uso de bosques aleatorios como el clasificador del método mejoró significativamente los resultados, en comparación a los clasificadores usuales como KNN y SVM. Esto se explica porque los bosques aleatorios se caracterizan por ser más robustos ante el desbalance de clases, ya que las decisiones que se toman internamente en los nodos generan que el modelo se especialice en identificar características propias de cada clase. De esta forma, la predicción de las clases es más flexible, en comparación con el algoritmo utilizado por

SVM. Si bien, el SVM es robusto para clasificaciones binarias, estos incorporan un concepto de distancia que se relaciona con el hiperplano que divide las dos clases; y esta distancia tiende a clasificar los datos beneficiando la clase que posee mayor cantidad de observaciones [3]. Por estos motivos, los bosques aleatorios generaron un mejor ACA, considerando que nuestro problema estaba significativamente desbalanceado (77.5 % - 22.5 %), como se mencionó en la sección de la base de datos, y que representa la distribución epidemiológica de la neumonía en la vida real.



Figura 7: Extracción insatisfactoria del pulmón de una de las radiografías.

El ACA obtenido en la etapa de clasificación de bounding boxes fue considerablemente alto (85.14 %), sin embargo, no es un resultado que se acerque a caracterizar un método más robusto. Si bien, se espera que esta etapa arroje resultados más cercanos al 100 % antes de pasar a la clasificación de la imagen completa, se decidió hacer esta transición para evaluar cómo era el comportamiento del problema en un escenario más realista, es decir, teniendo las imágenes de radiografías completas. A pesar de esto, los resultados fueron satisfactorios, pues el valor del ACA disminuyó únicamente en un 14.72 %. Dicha disminución del ACA está relacionada a la calidad de la extracción del pulmón, ya que no se pudo evaluar la calidad de la extracción del pulmón con parámetros como el índice de Jaccard, es decir, no se contó con anotaciones, en este caso máscaras. En consecuencia, no se pudo realizar mejoras significativas en el algoritmo de segmentación y extracción del pulmón. Si se hubiese contado con un indicador de la efectividad de la segmentación se podría haber implementado mejoras en el algoritmo de segmentación ya que la evaluación realizada fue cualitativa. Un ejemplo de una de las extracciones inadecuadas se muestra en la Figura 7. Cabe resaltar que la eficiencia del algoritmo está relacionada con la calidad de

la base de datos. Por ejemplo, en la Figura 8, se puede ver que la imagen de la radiografía no se encuentra centrada ni alineada. Esto genera que la segmentación del pulmón sea errónea además de generar una máscara poco precisa.

Referencias

- [1] RSNA Pneumonia Detection Challenge: Kernels, 2018. Recuperado el 6 de Abril de <https://www.kaggle.com/c/rsna-pneumonia-detection-challenge/kernels>. 2
- [2] N. Center. Algorithm better at diagnosing pneumonia than radiologists, 2019. Recuperado el 1 de Marzo de 2019 de: <https://med.stanford.edu/news/all-news/2017/11.html>. 1
- [3] A. Criminisi and J. Shotton, editors. *Decision Forests for Computer Vision and Medical Image Analysis*. Springer London, 2013. 7
- [4] O. M. de la Salud. Neumonía: Datos y cifras, n.d. Recuperado el 1 de Marzo de 2019 de: <https://goo.gl/YdMuyi>. 1
- [5] A. Depeursinge, A. S. Chin, A. N. Leung, D. Terrone, M. Bristow, G. Rosen, and D. L. Rubin. Automated classification of usual interstitial pneumonia using regional volumetric texture analysis in high-resolution computed tomography. *Investigative Radiology*, 50(4):261–267, Apr. 2015. 6
- [6] Merkow.J. Pneumonia detection in chest radiographs. Recuperado el 6 de Abril de : <https://arxiv.org/pdf/1811.08939.pdf>, year =2018. 2, 5
- [7] R. S. of North America. Rsn pneumonia detection challenge, 2018. Recuperado el 1 de Marzo de 2019 de: <https://www.kaggle.com/c/rsna-pneumonia-detection-challenge>. 5
- [8] P. Rui. National ambulatory medical care survey, 2015. Recuperado el 1 de Marzo de : https://www.cdc.gov/nchs/data/nhamcs/web_tables/2015_ed_web_tables.pdf. 1
- [9] M. Saad, Z. Muda, N. Ashaari, and H. Hamid. Image segmentation for lung region in chest x-ray images using edge detection and morphology. *IEE International Conference on Control System, Computing and Engineering*, pages 46–50, 2014. 2
- [10] M. N. Saad, Z. Muda, N. S. Ashaari, and H. A. Hamid. Image segmentation for lung region in chest x-ray images using edge detection and morphology. In *2014 IEEE International Conference on Control System, Computing and Engineering (ICCSC 2014)*, pages 46–51, Nov 2014. 6
- [11] P. N. R. Shabnam and S. M. Mohamed. Detection of pneumonia in chest x-ray images. *Journal of X-Ray Science and Technology*, 19(4):423–428, 2011. 6
- [12] A. Sharma, D. Raju, and S. Ranjan. Detection of pneumonia clouds in chest x-ray using image processing approach. *IEE journal*, pages 23–28, 2013. 2

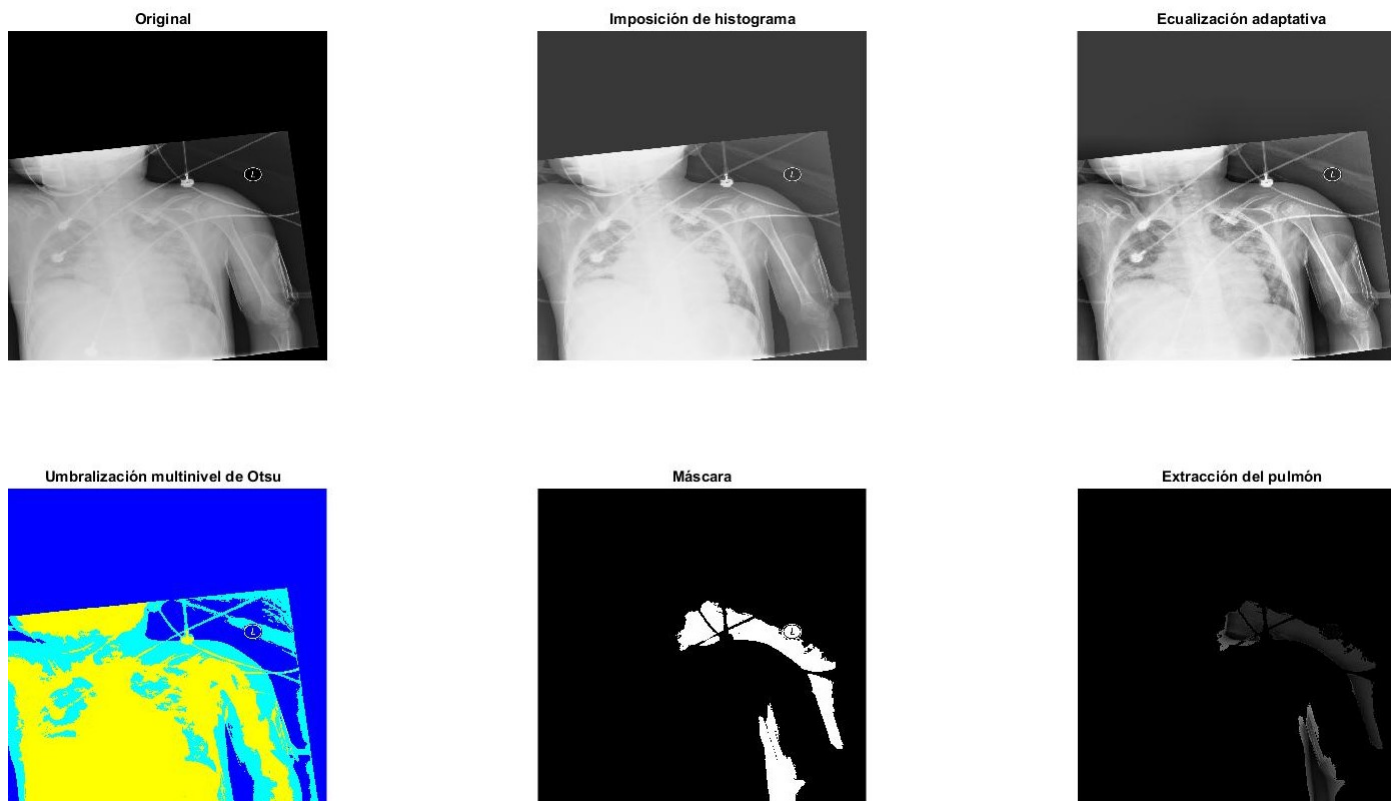


Figura 8: Resultados de las operaciones sobre las imágenes. Se puede observar una mala extracción de los pulmones debido a la pobre calidad inicial de la imagen de la radiografía

6.. Créditos

- Alejandra Barrera Suárez: Experimentos con histogramas de gradientes orientados, experimentos con textones, artículo final, algoritmo para clasificación a partir de bounding boxes.
- Jessica Castillo Rojas: Generación de documentos entregables (demo y evaluación final), artículo final, algoritmo de clasificación para imagen completa de radiografía, algoritmo para clasificación a partir de bounding boxes.
- Juan David García: Balance de la base de datos, algoritmo para bosques aleatorios, artículo final, algoritmo para clasificación a partir de bounding boxes.

Certificamos que Alejandra Barrera, Jessica Castillo y Juan David García tuvieron una contribución equivalente al proyecto.

Nota: Este documento fue realizado en \LaTeX