

Segmentación de Enfermedades Pulmonares Intersticiales en Imágenes de TCAR mediante Análisis de Textura y una Red Convolutiva tipo U-Net

C.G. Colin Tenorio^{1*}, A.R. Mejía Rodríguez¹, G. Dorantes Méndez¹

¹ Facultad de Ciencias, Universidad Autónoma de San Luis Potosí, San Luis Potosí, México.
**carmen.colin@alumnos.uaslp.edu.mx*

Resumen

En este trabajo se presenta una estrategia para poder realizar la segmentación de Enfermedades Pulmonares Intersticiales (EPI), específicamente Fibrosis Pulmonar Idiopática (FPI) y Enfisema Pulmonar (EP) de forma simultánea, en imágenes de Tomografía Computarizada de Alta Resolución (TCAR). Se realizó un procesamiento mediante análisis de textura para resaltar las características que diferencian a cada una de las enfermedades, utilizando estadísticos de primer y segundo orden. A partir de las imágenes procesadas se entrenó una red convolutiva U-Net, utilizando solo 10 rebanadas de un subvolumen de datos. Se propusieron 3 estrategias de entrenamiento, combinando la imagen de intensidad y diferentes estadísticos de textura, como desviación estándar, energía y homogeneidad. Se obtuvieron las segmentaciones de las diferentes EPI utilizando 10 redes neuronales entrenadas de distinta forma. Los resultados obtenidos mostraron valores de exactitud de hasta un 87% para la clase de FPI, y un valor de índice de similitud DICE e índice de intersección sobre la unión (IoU) de 75% y 78% respectivamente, resultados comparables a lo establecido en la literatura pero utilizando poca información en el entrenamiento de la red.

Palabras clave: Análisis de textura, Enfermedades pulmonares intersticiales, Red neuronal U-Net, Segmentación.

1 Introducción

Las enfermedades pulmonares intersticiales (EPI) son un grupo de trastornos crónicos que involucran todo el parénquima pulmonar y el intersticio alveolar, que causan diferentes grados de inflamación y fibrosis, que al progresar afecta la capacidad pulmonar. La Fibrosis Pulmonar Idiopática (FPI), representa poco más de la mitad (55%) de los casos de EPI, con una supervivencia media de 3 a 6 años, además de ser una enfermedad de difícil diagnóstico y con frecuencia requiere la colaboración de un grupo de médicos expertos en EPI para lograr tener un diagnóstico en consenso. La tomografía computarizada de alta resolución (TCAR) es una de las herramientas más utilizadas para la detección de las EPI. Sin embargo, la detección manual por el médico experto puede llegar a ser una tarea difícil, debido a que algunas veces la imagen tiene pocos detalles y los patrones de la enfermedad son poco perceptibles, particularmente en etapas tempranas de la enfermedad [1].

Por otra parte, el análisis computacional mediante redes neuronales convolucionales (CNN) ha permitido la segmentación de imágenes médicas con muy buenos resultados [2, 3]. Trabajos anteriores han utilizado CNN's para la segmentación semántica donde a diferencia de la clasificación, que predice una clase para toda la entrada, se requiere predecir una clase para cada píxel o voxel de la imagen de entrada. Las CNN's, son una excelente herramienta para segmentar imágenes médicas debido a que logran percibir detalles finos de la imagen, pueden definir superficies de decisión no lineales, y tienen procedimientos de entrenamientos automáticos. Sin embargo, requieren una gran cantidad de datos y

una gran capacidad computacional para procesarlos. En la literatura se encuentra una red neuronal denominada U-Net, la cual es una red neuronal totalmente convolucional (FCNN) que puede funcionar con pocos datos de entrenamiento. Este tipo de red ha sido probada para la segmentación de imágenes de microscopía electrónica de transmisión con buenos resultados [2], y por nuestro grupo de trabajo en imágenes de TCAR para la segmentación de EPI como la FPI y el EP utilizando solo la información de intensidad de las imágenes, obteniendo un porcentaje de exactitud de 87 %, resultado comparable con la delimitación manual de un experto [3].

Por otro lado, el análisis de textura permite extraer y resaltar las características de textura en una imagen para poder clasificar regiones de interés, proporciona una evaluación objetiva y cuantitativa de la heterogeneidad de un objeto en la imagen, analizando la distribución y relación de niveles de gris de píxeles o voxeles. Se han utilizado algunas métricas de textura de primer y segundo orden, como la varianza, entropía y desviación estándar para la segmentación de diversas enfermedades [4, 5]. Sin embargo, en ocasiones una métrica de textura solo puede ser útil para una enfermedad y no para varias enfermedades en una misma imagen, debido a ésto se considera que la combinación de diversas métricas de textura, de primer o segundo orden, puede ser una buena estrategia para poder determinar varias enfermedades al mismo tiempo. El procesamiento de imágenes mediante métricas de textura permite resaltar las características, patrones y detalles en la imagen para poder distinguir mejor entre la región de tejido sano y la región de tejido enfermo.

Por tal motivo, el objetivo de este trabajo es proponer una estrategia que permita segmentar las diferentes EPI en una imagen de TCAR de manera automática, con el fin de evitar la subjetividad del diagnóstico, además de detectar la enfermedad en una etapa temprana. La estrategia propuesta consiste en combinar las dos herramientas que ya han sido utilizadas en trabajos anteriores, el uso de la red neuronal tipo U-Net y el análisis de textura en las imágenes.

2 Metodología

2.1 Preprocesamiento de las Imágenes Médicas

Para el desarrollo de este trabajo se utilizó información de un paciente diagnosticado con EPI, FPI y Enfisema Pulmonar (EP), que consiste en imágenes de TCAR proporcionadas por el Instituto Nacional de Enfermedades Respiratorias (INER). El volumen está conformado por 55 cortes de TCAR con un tamaño de 512×512 voxeles y un tamaño de voxel de $0.63 \times 0.63 \times 9.00$.

Para este trabajo, se seleccionó un subvolumen representativo conformado por 10 cortes de la base de datos, 3 denominados datos de entrenamiento y los restantes datos de prueba, con el objetivo de usar solo un porcentaje de nuestro volumen de datos y evaluar la red con la mayor cantidad de imágenes. Es importante mencionar que los datos de entrenamiento fueron seleccionados de manera que tuvieran información de las dos enfermedades pulmonares así como de los rasgos más característicos de cada una de ellas. Las imágenes fueron preprocesadas mediante una delimitación manual de las secciones de interés, y se etiquetaron en 4 clases diferentes: Enfisema pulmonar, Fibrosis pulmonar idiopática, Tejido Sano y Fondo. Las segmentaciones fueron revisadas y validadas por un experto clínico del INER [3].

2.2 Análisis de Textura

El análisis de textura permite evaluar la distribución de intensidades y la organización espacial en una imagen. Las características de textura se pueden calcular mediante diferentes métodos, pero en este trabajo nos concentraremos en los métodos estadísticos, los cuales analizan la distribución espacial de los niveles de gris, mediante el cálculo del histograma y las matrices de textura de segundo y mayor orden. Dependiendo del número de píxeles que definen la característica local se pueden clasificar en estadísticos de primer orden (un píxel), de segundo orden (dos píxeles) y de orden superior (tres o más píxeles) [6].

2.2.1 Estadísticos de Primer Orden

El enfoque de histograma para el análisis de textura se basa en las concentraciones de valor de intensidad en toda o parte de una imagen representada como un histograma, el cual contiene la información estadística de primer orden sobre la imagen.

En este trabajo se seleccionaron dos medidas de primer orden, la desviación estándar y la entropía. De acuerdo a las características de las EPI, se propuso que éstas medidas podrían resaltar mejor la información en las imágenes de TCAR. La desviación estándar (DE) muestra cuanta variación o dispersión existe respecto del promedio (valor medio o esperado).

Una DE baja indica que los puntos de datos tienden a ser muy cercanos a la media, mientras que la DE alta indica que los puntos de datos se extienden en un amplio rango de valores. La DE matemáticamente viene dada por:

$$S = \sqrt{\frac{\sum_{i,j}^n (X_{i,j} - \bar{X})^2}{n - 1}} \quad (1)$$

donde \bar{X} representa la media, $X_{i,j}$ cada uno de los niveles de intensidad de los píxeles en la imagen, y n el número de píxeles.

Por otro lado, la entropía es una medida estadística de la aleatoriedad que se puede utilizar para caracterizar la textura de la imagen de entrada. Se espera que una imagen suave tenga una baja entropía y alta entropía será para aquellas con mayor textura. La entropía se define como:

$$E = \sum_{i,j}^{n-1} -P_{i,j} \ln(P_{i,j}) \quad (2)$$

donde $P_{i,j}$ representa la probabilidad de ocurrencia del píxel en el renglón i de la columna j de la matriz.

2.2.2 Estadísticos de Segundo Orden

El método *Gray Level Cooccurrence Matrix* (GLCM) es una forma de extraer características de textura estadísticas de segundo orden [7]. Una GLCM es una matriz donde el número de filas y columnas es igual al número de niveles de grises en la imagen. Haralick [8], sugirió el uso de GLCM para el análisis de textura de imágenes, este método está basado en las distribuciones de probabilidad conjunta de pares de píxeles.

El elemento de matriz $P(i, j|d, \theta)$ contiene valores de probabilidad estadística de segundo orden para los cambios entre los niveles de gris i y j a un desplazamiento particular $d=($ distancia en renglones, distancia en columnas), en un ángulo particular (θ°). La relación espacial entre el píxel de referencia y su vecino puede ser a 0° , 45° , 90° y 135° . Es decir, que cada elemento $P(i, j)$ indica cuántas veces aparece el valor i junto al valor j , en la relación espacial dada. Así, de la GLCM, pueden extraerse diferentes métricas de textura, en este trabajo, se seleccionaron 4 de ellas: Correlación, Homogeneidad, Energía y Contraste. Las cuales se describen a continuación:

$$CONTRASTE = \sum_{i,j} |i - j|^2 P(i, j) \quad (3)$$

$$CORRELACION = \sum_{i,j} \frac{(i - \mu_i)(j - \mu_j) P(i, j)}{\sigma_i \sigma_j} \quad (4)$$

$$ENERGIA = \sum_{i,j} P(i, j)^2 \quad (5)$$

$$HOMOGENEIDAD = \sum_{i,j} \frac{P(i, j)}{1 + |i - j|} \quad (6)$$

donde μ_i , μ_j , σ_i , y σ_j son las medias y desviaciones estándar de los valores de gris i y j . Siendo $P(i, j)$ la probabilidad de co-ocurrencia para un ángulo particular θ° .

2.3 Red Convolutiva U-Net

Una red neuronal convolutiva que consiste solo en capas convolucionales se conoce como FCN (*Fully Convolutional Network*), se construye solo a partir de capas conectadas localmente, como convolución, pooling y upsampling y puede funcionar independientemente del tamaño de la imagen original. Las FCN's tienen dos partes: el downsampling, donde se extrae la información, y la ruta de upsampling, donde se recupera la información espacial y al final se obtiene un mapa de segmentación. La red neuronal U-Net es una FCN encontrada en la literatura que utiliza muy pocos datos de entrenamiento y que además es recomendada para imágenes médicas por su gran capacidad para identificar detalles [2].

La arquitectura de la red U-Net utilizada en este trabajo consiste en un camino de contracción (submuestreo), uno de cuello de botella y otro de expansión (muestreo) [2, 3]. La ruta de contracción está compuesta de cuatro bloques, donde cada bloque está conformado por la aplicación de dos convoluciones de tamaño de filtro de 3×3 y después de cada convolución una unidad lineal rectificadora (ReLU), entre cada bloque se aplica una operación de agrupación máxima (max pooling) de tamaño 2×2 . En cada paso de reducción de muestreo se duplica el número de canales de características. Al inicio del primer bloque se tienen 64 canales de características, 128 en el segundo, y así sucesivamente. En la ruta de cuello de botella, se tienen dos capas convolucionales, cada convolución está seguida de una ReLU y una capa de Dropout, etapa de desactivación de neuronas de manera aleatoria, utilizada para evitar el sobreajuste.

La ruta expansiva está compuesta de cuatro bloques, donde cada bloque consiste en un muestreo ascendente del mapa de características y de una convolución de tamaño 2×2 (convolución ascendente) que reduce a la mitad el número de canales de características, una concatenación con el mapa de características recortado, correspondiente a la ruta de contracción, cada bloque está conformado por dos convoluciones de tamaño 3×3 , y cada convolución seguida de una ReLU. En la capa final, se utiliza una convolución 1×1 para asignar cada vector de características de 64 componentes al número deseado de clases [2].

2.4 Estrategias de Entrenamiento

Para el entrenamiento de las redes neuronales convolucionales tipo U-Net, se propusieron dos técnicas de entrenamiento, la primera utilizando imágenes correspondientes a una sola métrica de textura, y la segunda utilizando imágenes de dos métricas de textura. Antes de la etapa de entrenamiento, se procesaron las imágenes de prueba y de entrenamiento utilizando las métricas de textura. Se les aplicó un filtro de DE, un filtro de entropía y un filtro que calculaba la GLMC en cada una de las imágenes. Se realizaron pruebas con un kernel de 3×3 , 5×5 y 7×7 , el kernel de 5×5 presentó los mejores resultados ya que resaltó mejor los bordes entre las diferentes regiones de interés y se logró distinguir mejor entre las diferentes regiones. Es importante mencionar que al aplicar el filtro de entropía el resultado no logró mejorar la imagen original, por el contrario, la imagen tuvo menos detalles y el área de FPI no logra distinguirse visualmente. Por otra parte, en el caso de los filtros de contraste y correlación la imagen no logró mejorar los detalles en la imagen original, y las regiones de interés no se perciben ni se distinguen fácilmente. Por lo tanto, estas últimas métricas de textura se descartaron en el proceso de entrenamiento. La primera estrategia, consistió en seleccionar tres imágenes de entrenamiento con sus respectivas etiquetas, correspondientes a una métrica de textura. En total se ingresaron seis imágenes en este primer entrenamiento y se realizó para cada una de las métricas: Homogeneidad, DE y energía. El segundo entrenamiento, consistió en seleccionar 6 imágenes de entrenamiento, de igual manera con sus respectivas etiquetas, 3 de ellas correspondientes a una métrica de textura y las otras 3 imágenes a una métrica distinta. Las combinaciones de métricas fueron las siguientes: Intensidad de gris y energía, Intensidad de gris y DE, DE y energía, y por último Energía y homogeneidad.

Para ambas estrategias, se definió un tamaño de parche de 256×256 píxeles, 5 el tamaño de cada mini lote, y se tomaron 50 lotes por cada imagen, puesto que trabajos anteriores [3], mencionan que esos valores funcionan de manera adecuada para este tipo de entrenamiento e imágenes. Por lo tanto, en cada iteración de la época se extraían 150 o 300 mini lotes, dependiendo de la estrategia de entrenamiento con 3 y 6 imágenes de entrada, donde cada mini lote contenía 5 parches de tamaño 256×256 (5 de la imagen original y 5 de las etiquetas correspondientes). Es decir, en cada una de las iteraciones

se generaban 750 o 1500 parches de las imágenes de entrada. Los parches se extraen en posiciones aleatorias en las imágenes, y se les aplica una transformación aleatoria con el objetivo de aumentar los datos (*data augmentation*), la transformación puede ser una rotación aplicada en 4 diferentes ángulos (30, 45, 60 y 90 grados), o una reflexión en dirección al eje X o Y. El número de épocas fue establecido en 10, por tanto, la primer estrategia de entrenamiento consistió en 1500 iteraciones y la segunda de 3000 iteraciones. Finalmente, las propuestas de entrenamiento de las redes neuronales se evaluaron con las imágenes de prueba.

3 Resultados y Discusión

En la Fig. 1 se puede apreciar la imagen original de intensidad de gris así como el resultado de la aplicación de diferentes filtros de textura y se puede observar que cada una de las imágenes obtenidas presenta nueva información de las EPI, además de que la textura nos permite tener una nueva interpretación de la imagen.

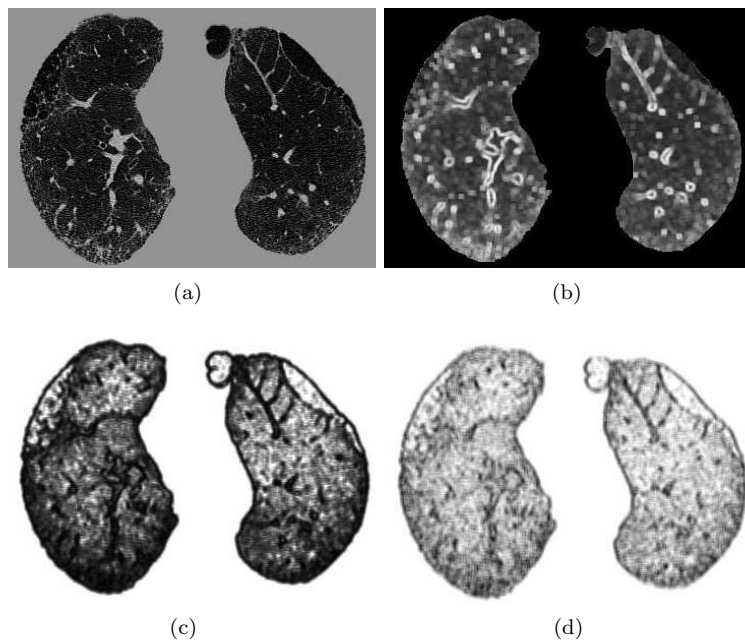


Fig. 1. Procesamiento de las imágenes mediante el uso de análisis de textura: a) Imagen original de intensidad de gris, b) resultado de aplicar el filtro de DE, c) resultado del filtro de energía, y d) resultado del filtro de homogeneidad; en todos los casos se utilizó un kernel de 5×5 .

3.1 Redes entrenadas con una métrica de textura

En la primera estrategia de entrenamiento se seleccionaron tres métricas de textura: DE, energía y homogeneidad. Con las imágenes de cada una de las métricas de textura y las imágenes de intensidad originales se entrenaron las redes neuronales tipo U-Net, como resultado, se obtuvieron las segmentaciones de cada una de las regiones de interés: FPI, EP, tejido sano (TS) y fondo (FO). En la Fig. 2, se muestra un ejemplo de la segmentación obtenida de la red entrenada con imágenes de DE y la segmentación manual de las clases, usada como referencia.

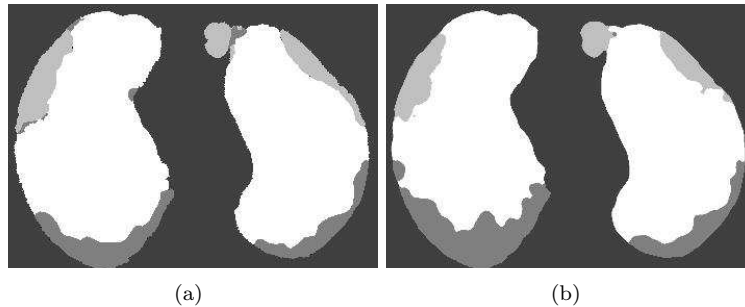


Fig. 2. Ejemplo del resultado de la segmentación de las redes entrenadas: a) Segmentación manual de las diferentes clases y b) Resultado de la segmentación de la red entrenada con imágenes de DE, donde blanco corresponde al tejido sano, gris claro a EP, gris medio a FIP y gris oscuro al fondo.

Para hacer un análisis cuantitativo de cada una de las redes neuronales y poder tener una comparativa, se calcularon los índices de desempeño: el índice de similitud DICE, el cual indica que tanto se traslapan las estructuras presentes en ambas segmentaciones y puede tener un valor entre cero, lo que significa una superposición nula y uno, que significa una superposición perfecta, la exactitud y el índice de intersección sobre la unión (IoU), como se muestra en la tabla 1. En la red entrenada con las imágenes de intensidad de gris, como se puede apreciar, el valor de la exactitud en la clase de FPI y TS tiene valores por arriba de 87%. En el caso del índice de IoU, todas las clases presentan valores por arriba del 70% y el índice DICE tuvo valores muy aceptables para las clases de FPI y TS por arriba de 75%. En general, la red tiene un buen desempeño para la clasificación de los píxeles en su clase verdadera en la clasificación de FPI y TS.

Para la red entrenada con imágenes de energía se puede observar que la exactitud para la clase de TS tuvo un muy buen desempeño con un valor de 96%. En cuanto al índice de IoU, el valor para la clase de EP fue muy bajo con un 48%, el índice DICE, obtuvo valores arriba del 71% para la clase de TS y FPI. En general, la red desempeña un buen trabajo en las clases de FPI y TS.

Para la red neuronal entrenada con imágenes de DE, en el caso del índice de exactitud, la red presenta un valor por arriba del 81% para las clases de FPI y TS. De igual manera, el índice de IoU para las clases de FPI y EP, presenta valores muy aceptables por arriba de 70% y el índice DICE presenta resultados por arriba del 72% para las 3 clases. Esta red, en general, tuvo muy buenos resultados en todas las clases.

Por último, para la red neuronal entrenada con imágenes de homogeneidad, el índice de exactitud presenta un valor muy alto para la clase de TS con un 99%, sin embargo, para la clase de FPI presenta un valor poco aceptable con un 20%, lo cual representa un mal desempeño para esta clase. El índice de IoU presenta valores muy bajos para la clase de FPI con un valor de 20% y en el caso del DICE, tiene un valor muy favorable para la clase de TS con un 92%. En general, esta red tuvo resultados poco favorables para la clasificación de los píxeles correspondientes a FPI, sin embargo, tuvo un gran desempeño en la clasificación de los píxeles de TS. En particular, la red de DE y la red de intensidad de gris presentaron los valores más altos para los tres índices de desempeño en todas las clases de interés y se consideran las mejores redes neuronales en la primera estrategia de entrenamiento.

3.2 Redes entrenadas con dos métricas de textura

Debido a que en la sección anterior se observó que los resultados presentan un gran traslape entre las clases FPI y EP con la clase de TS, se propuso la segunda técnica de entrenamiento que consistió en entrenar la red neuronal con imágenes correspondientes a dos métricas distintas. Las combinaciones para el proceso de entrenamiento fueron: Intensidad-DE, Intensidad-Energía, DE-Homogeneidad y Energía-Homogeneidad. En general, al obtener las segmentaciones mediante esta estrategia se observó que los resultados fueron poco favorables y muy distintos a las etiquetas originales. Lo cual puede deberse a que

Tabla 1: Índices de desempeño para las redes neuronales entrenadas con una métrica de textura en cada una de las clases de interés.

| Métrica | Índice | Clases | | |
|---------------------|-----------|---------------|----------------|---------------|
| | | FPI | EP | TS |
| Intensidad de gris | DICE | 0.754 ± 0.056 | 0.699 ± 0.135 | 0.934 ± 0.015 |
| | Exactitud | 0.873 ± 0.052 | 0.736 ± 0.105 | 0.909 ± 0.032 |
| | IoU | 0.786 ± 0.034 | 0.712 ± 0.106 | 0.715 ± 0.055 |
| Energía | DICE | 0.711 ± 0.061 | 0.533 ± 0.383 | 0.937 ± 0.020 |
| | Exactitud | 0.661 ± 0.071 | 0.496 ± 0.373 | 0.967 ± 0.019 |
| | IoU | 0.642 ± 0.065 | 0.488 ± 0.367 | 0.568 ± 0.125 |
| Desviación Estándar | DICE | 0.728 ± 0.103 | 0.728 ± 0.140 | 0.935 ± 0.029 |
| | Exactitud | 0.815 ± 0.046 | 0.742 ± 0.1386 | 0.920 ± 0.047 |
| | IoU | 0.762 ± 0.058 | 0.704 ± 0.140 | 0.669 ± 0.096 |
| Homogeneidad | DICE | 0.331 ± 0.074 | 0.648 ± 0.244 | 0.921 ± 0.017 |
| | Exactitud | 0.207 ± 0.055 | 0.567 ± 0.255 | 0.991 ± 0.003 |
| | IoU | 0.206 ± 0.055 | 0.558 ± 0.251 | 0.461 ± 0.058 |

solo una de las dos métricas de entrenamiento tiene un buen desempeño en el proceso de entrenamiento y la otra métrica no. Por tanto, el orden de ingreso de las imágenes en la red puede que afecte el proceso de entrenamiento.

Para corroborar la idea anterior, se realizó una prueba adicional que consistió en modificar el orden en que se ingresaban las imágenes. De manera que anteriormente las imágenes se ingresaron como "aaabbb", donde *a* corresponde a una imagen de una métrica de textura y *b* a una imagen de entrada de otra métrica de textura, se modificó el orden de ingreso de las imágenes como "bbbaaa", y "ababab". Esta técnica se probó para las métricas de intensidad de gris y DE y se comprobó que el orden de ingreso de las imágenes es capaz de influir en el proceso de entrenamiento de la red, así como también en la segmentación resultante.

4 Conclusiones

En este trabajo, se presentó una nueva estrategia para poder segmentar las EPI en un subvolumen de TCAR, en la cual se combinó el uso de una red neuronal tipo U-Net y el uso del análisis de textura mediante estadísticos de primer y segundo orden. Como resultado del entrenamiento de 10 redes neuronales, se obtuvieron las segmentaciones de cada una de ellas para las clases de interés. Cada métrica de textura aporta diferente contribución lo cual se debe a que cada una extrae diferente información de la imagen, y se comprueba que al utilizar el mismo subvolumen de imágenes, el uso del entrenamiento con diferentes métricas de textura, modifica el resultado final de la segmentación. Los resultados de la primera estrategia comprobaron que la red tipo U-Net es una muy buena estrategia para segmentar imágenes médicas debido a que utiliza solo un pequeño subvolumen de la base de datos, y al ingresar

pocas imágenes de entrenamiento se pueden obtener resultados muy favorables. Adicionalmente, se observó que al combinar las métricas de textura, solo algunas imágenes de entrenamiento lograron contribuir a un entrenamiento eficiente y las otras imágenes de entrada no lograron tener una alta contribución. Por otro lado, al modificar el orden en que se ingresaban las imágenes los resultados fueron muy distintos y puede ser debido a la configuración y a la estructura de la red U-Net en la cual no se puede utilizar más de una métrica de textura, lo que indica que la red en Matlab está diseñada para ingresar imágenes de un solo tipo.

Como trabajo futuro se plantea modificar la estructura de la red U-Net para que una métrica de textura pueda tener mayor aportación en el proceso de entrenamiento que la otra métrica. También se propone modificar la red para que el orden de ingreso de las imágenes no afecte el proceso de entrenamiento. En este trabajo se exploraron solo algunas métricas de textura, sin embargo se propone buscar otras métricas de textura que también puedan extraer información relevante durante el proceso de entrenamiento. Además se propone incrementar el número de imágenes de prueba para evidenciar el desempeño de la estrategia propuesta y observar si se obtienen mejores resultados. Se probará también incrementar los datos (*data augmentation*) con mayor cantidad de transformaciones, como traslaciones y rotaciones distintas, en los parches para ampliar la cantidad de las imágenes de entrenamiento. Por último, se propone implementar la modificación de la red U-Net en el programa Python.

Declaración de conflictos de interés

Los autores declaran no tener ningún conflicto de interés para este trabajo.

Agradecimientos

A los Drs. Mayra Mejía, Sonia Charleston y Tomás Aljama por su colaboración en la obtención de los datos.

Referencias

- [1] M. Mejía, I. B. Roldán, and T. Mateos, “Primer Consenso Mexicano sobre Fibrosis Pulmonar Idiopática,” *Neumol Cir Torax*, vol. 75, pp. 32–51, Mar. 2020.
- [2] O. Ronneberger, P. Fischer, and T. Brox, “U-net: Convolutional Networks for Biomedical Image Segmentation,” in *MICCAI*, pp. 234–241, Springer International Publishing, 2015.
- [3] S. H. Juárez, A. R. M. Rodríguez, E. R. A. Santana, S. C. Villalobos, A. T. A. Corrales, *et al.*, “Semantic Segmentation of Lung Tissues in HRCT Images by Means of a U-Net Convolutional Network,” in *IFMBE*, pp. 426–434, Springer International Publishing, Oct. 2019.
- [4] K. Vaishnav and V. Malode, “New approach of statistical analysis for lung disease diagnosis using microscopy images,” in *ICACDOT*, pp. 378–383, Sept. 2016.
- [5] M. Punal, K. Nanditha, D. Vamsha, and B. Lakshmi, “Statistical analysis of lung CT images for ground glass diagnosis,” in *AEEICB*, pp. 347–350, Feb. 2016.
- [6] G. Srinivasan and G. Shobha, “Statistical Texture Analysis,” in *WASET*, vol. 2, pp. 4268–4273, 2008.
- [7] F. Albreghsen, “Statistical Texture Measures Computed from Gray Level Cooccurrence matrices,” *Image Processing Laboratory, Department of Informatics*, Jan. 1995.
- [8] R. M. Haralick, K. Shanmugam, and I. Dinstein, “Textural Features for Image Classification,” *IEEE Trans. Syst. Man Cybern.*, vol. SMC-3, pp. 610–621, Nov. 1973.