

Detección automática de tejido blando nasal en CT-Scan y MRI utilizando Random Forests

Pazos, B. A.^{1,2,3}, Navarro, P.^{1,2,3}, Morales, A. L.^{1,2,3}, Cintas, C.^{1,2,3}, Trujillo, A.^{1,2}, de Azevedo, S.¹, González-José, R., Delrieux, C.²

¹ Instituto Patagónico de Ciencias Sociales y Humanas (IPCSH). Centro Nacional Patagónico (CENPAT), Puerto Madryn, Chubut.

² Laboratorio de Ciencias de las Imágenes, Departamento de Ingeniería Eléctrica y Computadoras. Universidad Nacional del Sur, Bahía Blanca.

³ Universidad Nacional de la Patagonia San Juan Bosco, Facultad de Ingeniería, Trelew, Chubut.

Abstract. Este trabajo presenta un procedimiento sencillo para la detección automática de tejido blando nasal a partir de imágenes médicas de tipo CT-Scan y MRI. Para dichas imágenes, inicialmente se aplica un procedimiento de segmentación manual y técnicas de procesamiento digital de imágenes para la confección de un dataset de entrenamiento, conteniendo los valores de cada pixel de la imagen por un lado y 4 valores que representan el área de interés a detectar por otro. Posteriormente, se utilizan métodos de preprocesamiento de datos sobre estas imágenes para ser utilizadas como conjunto de entrenamiento, validación y prueba de un estimador, empleando un algoritmo basado en un modelo de regresión multi-salida denominado Random Forest. Este último es capaz de detectar automáticamente el área de la imagen en la cual se encuentra el tejido blando nasal prediciendo simultáneamente los 4 valores que representan los bordes del área de interés mencionada.

1 Introducción

En múltiples escenarios de segmentación y reconocimiento en imágenes resulta de gran importancia la determinación de un área a segmentar, también llamada región de interés o ROI (por sus siglas en inglés). Esta técnica se utiliza muy frecuentemente por ejemplo en la detección y clasificación de objetos utilizando imágenes satelitales [10, 2], en la detección de rostros para estudiar los diferentes gestos faciales [3] y en la segmentación automática de región mamaria para detección temprana de cáncer de pecho [15], entre otros. Este proceso disminuye la cantidad de cómputo necesario para lograr una reconstrucción/segmentación satisfactoria al reducir la cantidad de píxeles con la cual se debe trabajar. Esta idea es aún más relevante en imágenes médicas de tipo CT-Scan y/o MRI, en las cuales se debe lidiar con un gran volumen de imágenes o *slices* que representan una o varias estructuras anatómicas pertenecientes a un mismo paciente.

Este trabajo está orientado a la obtención automática de los parámetros que definen una “ventana” o subsección de la imagen original en la cual se encuentra

el tejido blando nasal y que forma parte de la estructura nasofaríngea de un paciente. En las secciones siguientes se describirá con detalle cómo está compuesto el conjunto de datos de entrada (las imágenes médicas), y los procedimientos llevados a cabo para conformar un dataset que servirá para alimentar un algoritmo de regresión que pueda predecir la ubicación de la ROI antes mencionada. Al mismo tiempo, se presentarán las métricas que evaluarán al algoritmo y definirán si es o no factible obtener resultados exitosos con las técnicas empleadas frente a datos desconocidos provenientes de nuevos conjuntos de imágenes.

2 Materiales y métodos

A continuación se describirá el conjunto de datos utilizados en este trabajo como así también las operaciones de preprocesamiento aplicadas. Luego se presentará el estimador seleccionado y la configuración de sus parámetros. Para el desarrollo de código se utilizó el lenguaje de programación Python [14] y librerías como Numpy [16], Matplotlib [8] y Scikit-learn [13].

2.1 Conjunto de datos

Se utilizó un total de 806 imágenes en plano axial provenientes de 10 pacientes distintos. La cantidad de imágenes por paciente varía de acuerdo al grosor de las slices, dependiendo del tipo de estudio o del nivel de detalle indicado por el/la profesional que solicita el estudio en un primer lugar. Las imágenes, se obtienen como archivos con la extensión *.dcm* por el estándar *Digital Imaging and Communications in Medicine* (DICOM) [12]. Se realiza un proceso de anonimización, abriendo el archivo DICOM y separando sus componentes para acceder al arreglo de píxeles que componen la imagen de interés. Se repitió este proceso para cada uno de los archivos obteniendo las 806 imágenes en escala de grises de 512×512 píxeles.

Con base en este conjunto de imágenes, se creó y segmentó manualmente una copia de cada una de ellas utilizando el editor de imágenes *GNU Image Manipulation Program* (GIMP) [1], creando una máscara binaria en la cual el color blanco representa el tejido blando nasal presente en la imagen original y el color negro representa el resto de la imagen, es decir, aquellos componentes que no son de interés para el objetivo de este trabajo. Este procedimiento puede apreciarse en la Fig. 1. Ambos conjuntos de imágenes fueron almacenados en carpetas diferentes respetando el nombre de los archivos de modo tal que fuera posible realizar la correspondencia entre la imagen original y la imagen segmentada.

Posteriormente, se creó un algoritmo sencillo que consiste en procesar la imagen segmentada, buscando la primera y la última fila en la cual aparecen los píxeles blancos, guardando esos valores y procediendo del mismo modo con las columnas. De este modo, por cada imagen de entrada se guardaron 4 parámetros que determinan la ventana de píxeles de la imagen original que contienen tejido blando nasal. A modo de ejemplo y tomando la figura anterior como base, se realizó el recorte de ambas imágenes y su resultado puede observarse en la Fig. 2.

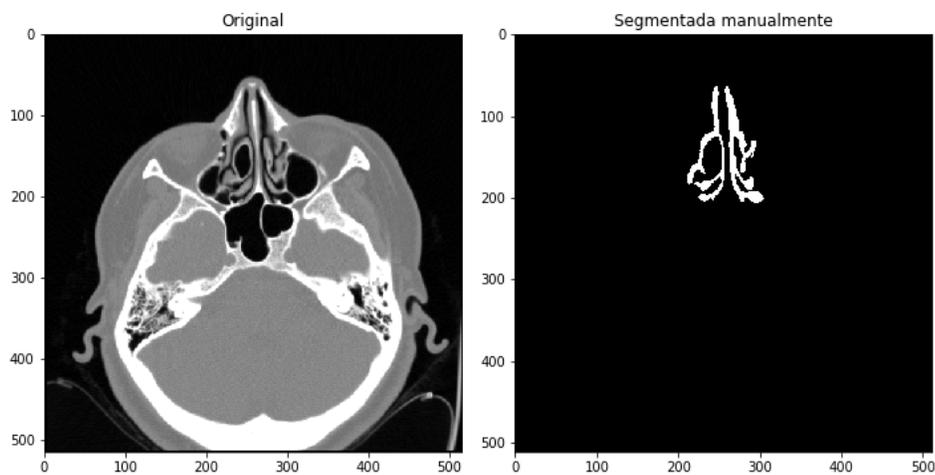


Fig. 1: Máscara binaria y segmentación de tejido blando nasal.

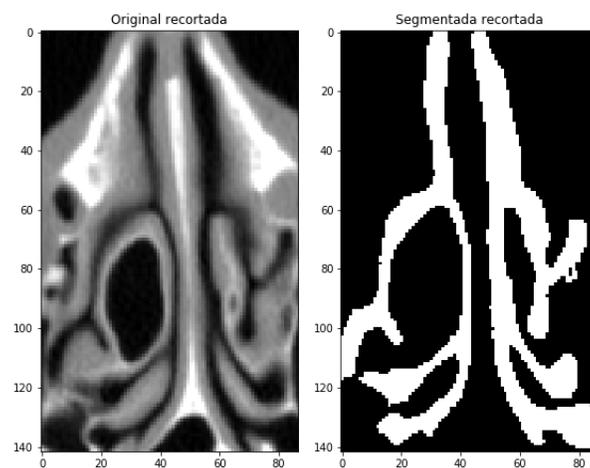


Fig. 2: Recorte de tejido blando nasal.

De este modo se conformó un dataset inicial que por un lado contiene los 512×512 píxeles de cada imagen original y por otro lado los 4 valores que permiten la obtención de la ventana de tejido blando nasal presente en cada una de las imágenes del primer conjunto.

2.2 Preprocesamiento de datos

Para reducir la complejidad intrínseca de los datos, se realizaron 2 operaciones de preprocesamiento:

- Normalización.
- Análisis de componentes principales (PCA, por sus siglas en inglés).

La operación de normalización permite evitar que el estimador utilizado tenga algún tipo de sesgo originado por las diferentes escalas en las que los datos puedan presentarse. Para ello, se utilizó el algoritmo *MinMaxScaler* que transforma los valores suministrados, llevándolos a una escala conocida, en nuestro caso entre 0 y 1. Mientras que para la segunda operación se utilizó el algoritmo *PCA* para reducir la dimensionalidad de los datos y en consecuencia, la complejidad del modelo en niveles generales.

El dataset final quedó conformado por los 15 PCs más significativos de cada imagen original como entradas (representando el 71% de la variabilidad) y como salida, por los valores que conforman la primera y última fila y columna de la imagen en la cual aparece un pixel de tejido blando nasal por cada una de las imagenes de entrada. Para lograr la automatización en la detección del área de interés y evitar tener que segmentar manualmente cada imagen médica que desee procesarse en un futuro, se procedió a utilizar técnicas de Machine Learning y dada la naturaleza del problema, se aplicó un algoritmo de regresión que permite predecir cada uno de los 4 valores que determinan la ubicación del tejido blando nasal.

2.3 Estimador elegido y configuración de sus parámetros

Como modelo de regresión se utilizó un algoritmo de regresión basado en Random Forest [4] debido a su capacidad de realizar predicciones de múltiples salidas (multiple output) para una misma entrada y por demostrar una gran versatilidad para problemas similares como los descritos en [7, 6, 5, 11]. Para el estimador utilizado, se configuraron sus parámetros como sigue:

- *criterion* = 'mae'
- *random_state* = 0
- *max_depth* = 20
- *n_estimators* = 200
- *n_jobs* = -1,

donde los valores de profundidad máxima y el número de estimadores fueron obtenidos mediante la utilización del algoritmo *Grid Search* que prueba para un mismo conjunto de datos, la combinación de diferentes parámetros de configuración del estimador y en base a los mejores resultados, determina los parámetros adecuados. Para la profundidad máxima se pusieron a prueba los valores: 2, 3, 4, 5, 10, 20 y 30, mientras que para la cantidad de estimadores se utilizaron los siguientes valores: 120, 140, 160, 180, 200.

Finalmente Grid Search obtuvo sus mejores resultados con 20 y 200, respectivamente. El parámetro *n_jobs* es igual a -1 dada la gran cantidad de datos de entrada y la necesidad de obtener la mejor performance posible del algoritmo utilizado, permite la utilización de todos los cores del CPU que se encuentren

Table 1: Resultados de entrenamiento.

Corte	R^2
1	0,940
2	0,928
3	0,923
4	0,945
5	0,952
Promedio	0,938
Desvío Estándar (+/-)	0,021

disponibles. Por último, para Grid Search se determinó que la cantidad de validaciones cruzadas sería igual a 5 por cada entrenamiento. En este caso, con 200 estimadores de 20 niveles de profundidad cada uno, el algoritmo alcanza su mejor performance con un R^2 promedio de 94,55%.

3 Resultados experimentales

Teniendo en cuenta el objetivo de este trabajo, el cual es poder introducir al algoritmo utilizado los datos provenientes de una imagen completamente nueva para este y obtener como resultado 4 valores que se correspondan el área en la cual se encuentra el tejido blando nasal de la imagen de entrada, se procedió a realizar la etapa de entrenamiento y los posteriores experimentos.

El conjunto de datos de entrada fue dividido en primera instancia con 600 imágenes utilizando el 75% de ellas para entrenamiento y el 25% para validación, dejando 206 imágenes para realizar pruebas con imágenes desconocidas para el algoritmo. Para entrenar y validar el estimador, se utilizó el algoritmo *shuffle split* creando 5 cortes diferentes del conjunto de entrenamiento. Luego, se procedió a validar el estimador utilizando R^2 como métrica, obteniendo un promedio de 93,83% con un desvío estándar de $\pm 2,12\%$ para los 4 valores predichos por el modelo, como indican los valores de la Tabla 1. Durante el período de prueba y utilizando las 206 imágenes restantes, se obtuvieron los siguientes resultados:

- Variancia explicada: 89,57%.
- R^2 promedio: 93,20% con un desvío estándar de $\pm 6,28\%$.
- Error absoluto promedio: 4,79 con un desvío estándar de $\pm 3,00$.

Estos resultados pueden observarse gráficamente en la Fig. 3, en la que por un lado se realizó el corte manualmente utilizando el algoritmo implementado por los autores de este trabajo sobre una de las 206 imágenes desconocidas para el estimador y por otro lado, el recorte realizado utilizando los parámetros predichos por tal estimador.

Para cuantificar la pérdida de datos, se utilizó la métrica *Intersection over Union* o (IoU) [9] que, como su nombre lo indica, es el cociente resultante de la división entre la intersección sobre la unión de las áreas cubiertas tanto por el recorte original como por el recorte obtenido producto del uso del algoritmo

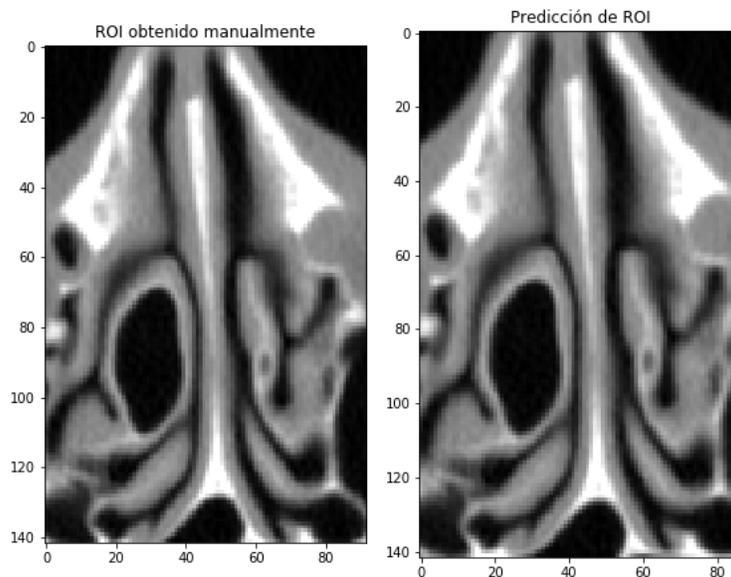


Fig. 3: ROI obtenido manualmente vs. predicción de ROI

estimador, permitiéndonos medir la similitud entre conjuntos de datos y se define como sigue:

$$J(A, B) = \frac{|A \cap B|}{|A \cup B|} = \frac{|A \cap B|}{|A| + |B| - |A \cap B|} ,$$

donde: $0 \leq J(A, B) \leq 1$.

En este caso, se obtuvo un IoU de 0.9202 (o 92,02% en términos porcentuales), lo cual representa una pérdida de un 7,98% de los datos.

4 Conclusión y trabajos futuros

Se presentó un modelo para la detección automática del área de tejido blando nasal presente en imágenes médicas de tipo CT-Scan y MRI utilizando técnicas de procesamiento digital de imágenes y Machine Learning obteniendo resultados satisfactorios. A futuro se plantea utilizar el modelo entrenado para ampliar el dataset inicial y permitir la segmentación automática del tejido blando nasal con vistas a un modelo de reconstrucción automática en 3D para la utilización en simulación de procedimientos quirúrgicos, estudiar problemas de respiración causados por algún tipo de obstrucción o lesión en la estructura anatómica, entre otras posibles aplicaciones.

References

1. GNU Image Manipulation Program. <https://www.gimp.org/>, [Online]

2. Abraham, L.: Vehicle detection and classification from high resolution satellite images <https://www.isprs-ann-photogramm-remote-sens-spatial-inf-sci.net/II-1/1/2014/isprsannals-II-1-1-2014.pdf>
3. Arora, V.: ROI segmentation for feature extraction from human facial images <https://arxiv.org/pdf/1207.2922.pdf>
4. Breiman, L.: Random Forests. *Machine Learning* 45(1), 5–32 (2001), <http://link.springer.com/10.1023/A:1010933404324>
5. Cootes, T.F., Ionita, M.C., Lindner, C., Sauer, P.: Robust and Accurate Shape Model Fitting Using Random Forest Regression Voting. pp. 278–291. Springer, Berlin, Heidelberg (2012), http://link.springer.com/10.1007/978-3-642-33786-4_21
6. Criminisi, A., Robertson, D., Konukoglu, E., Shotton, J., Pathak, S., White, S., Siddiqui, K.: Regression forests for efficient anatomy detection and localization in computed tomography scans. *Medical image analysis* 17(8), 1293–303 (dec 2013), <http://www.ncbi.nlm.nih.gov/pubmed/23410511>
7. Cuingnet, R., Prevost, R., Lesage, D., Cohen, L.D., Mory, B., Ardon, R.: Automatic Detection and Segmentation of Kidneys in 3D CT Images Using Random Forests. pp. 66–74. Springer, Berlin, Heidelberg (2012), http://link.springer.com/10.1007/978-3-642-33454-2_9
8. Hunter., J.D.: Matplotlib: A 2d graphics environment (2007), <https://aip.scitation.org/doi/abs/10.1109/MCSE.2007.55>
9. Jaccard, P.: Étude comparative de la distribution florale dans une portion des Alpes et des Jura. *Bulletin del la Société Vaudoise des Sciences Naturelles* 37, 547–579 (1901)
10. Kiadtikornthaweeeyot, W., Tatnall, A.R.L.: Automatic detection of a region of interest in a satellite image. http://a-a-r-s.org/acrs/administrator/components/com_jresearch/files/publications/TH2-3-1.pdf
11. Lindner, C., Thiagarajah, S., Wilkinson, J., Consortium, T., Wallis, G., Cootes, T.: Fully Automatic Segmentation of the Proximal Femur Using Random Forest Regression Voting. *IEEE Transactions on Medical Imaging* 32(8), 1462–1472 (aug 2013), <http://ieeexplore.ieee.org/document/6497663/>
12. Parisot, C.: The DICOM standard. *The International Journal of Cardiac Imaging* 11(3), 171–177 (sep 1995), <http://link.springer.com/10.1007/BF01143137>
13. Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., Vanderplas, J., Passos, A., Cournapeau, D., Brucher, M., Perrot, M., Duchesnay, E.: Scikit-learn: Machine learning in Python. *Journal of Machine Learning Research* 12, 2825–2830 (2011)
14. Rossum, Guido: Python reference manual (1995), <https://dl.acm.org/citation.cfm?id=869369>
15. Sahakyan, A., Sarukhanyan, H.: Segmentation of the Breast Region in Digital Mammograms and Detection of Masses. *IJACSA) International Journal of Advanced Computer Science and Applications* 3(2) (2012), www.ijacsa.thesai.org
16. Travis E, O.: A guide to NumPy (2006)