

Segmentación semántica de imágenes con sobreexpresión de la proteína HER2 utilizando Fuzzy Decision Trees

INF-575 Fuzzy Logic

Sebastián Ignacio Bórquez González
Departamento de Informática
Universidad Técnica Federico Santa María
Valparaíso, Chile
sebastian.borquez@sansano.usm.cl

Resumen—La proteína *human epidermal growth factor receptor 2* (HER2) es un biomarcador utilizado en muestras histológicas para el diagnóstico de distintos tipos de cáncer. La clasificación de los tejidos con sobreexpresión de la proteína HER2, es una tarea crucial realizada por los patólogos, que permite determinar si el paciente es apto para recibir la terapia dirigida a HER2. Sin embargo, este análisis es semi-cuantitativo, y utiliza características de las estructuras celulares expresado en variables lingüísticas. Avances en la digitalización de las muestras, han permitido el desarrollo de herramientas computacionales para asistir en el proceso de evaluación. El análisis computacional de estas imágenes no es una tarea trivial, distintos factores durante la preparación de las muestras, como el equipo utilizado para digitalizar la muestra, pueden resultar en ruido en la imagen. El siguiente trabajo propone un algoritmo que basado en *fuzzy logic* para en la segmentación semántica de estructuras celulares en imágenes de tejido, considerando la inclusión de conocimiento del dominio y el desbalance de clases.

I. DEFINICIÓN DEL PROBLEMA

La evaluación de la sobreexpresión de la proteína *human epidermal growth factor receptor 2* (HER2) presente en muestras histológicas de pacientes con cáncer, es un procedimiento clínico crucial realizado por los patólogos, que permite determinar el pronóstico de la enfermedad, junto con el tratamiento clínico. [1] La proteína HER2 se expresa normalmente en la superficie de la membrana celular del epitelio de distintos órganos, como pulmón, vejiga, páncreas, mama, estómago y próstata. Cuando la proteína HER2 está presente en cantidades normales, esta juega un papel importante en el desarrollo normal del crecimiento celular. Sin embargo, cuando hay una producción excesiva de la proteína HER2, es decir, existe sobreexpresión, las células se multiplican y crecen más rápidamente de lo normal, promoviendo la carcinogénesis en diversos tipos de cáncer [2].

En el caso del cáncer de mama, el análisis de la sobreexpresión de HER2 es fundamental para determinar cuando el paciente es apto para recibir la terapia dirigida a HER2, lo que incluye tratamientos con anticuerpos anti-HER2. Estudios indican que alrededor del 20% de los pacientes con cáncer de mama son HER2 positivos, es decir, presentan sobreexpresión

de la proteína HER2 [1]. El uso de la terapia dirigida HER2 junto con quimioterapia aumentan la supervivencia general solo de las personas que padecen cáncer de mama HER2 positivo. [3]

La inmunohistoquímica (IHC, de sus siglas en inglés *immunohistochemistry*) es el procedimiento histológico comúnmente usado para analizar la sobreexpresión del biomarcador HER2. En forma general, la IHC es una técnica que permite detectar e identificar anticuerpos específicos en células o muestras de tejido, basándose en el principio de unión existente entre antígenos y anticuerpos objetivos. En el caso de la evaluación de sobreexpresión de HER2, la proteína HER2 corresponde al antígeno objetivo [4]. La hematoxilina y la diaminobencidina (DAB) son compuestos químicos usados durante el procedimiento de IHC para visualizar los núcleos y las membranas celulares. Los efectos de aplicar estos compuestos a los tejidos mediante la técnica IHC puede apreciarse en la Figura 1. La hematoxilina se usa como contratinción que produce tinción azul-violeta de los elementos nucleares. DAB produce una tinción marrón oscuro en la membrana celular solo cuando hay sobreexpresión de HER2.

Las muestras tratadas con IHC, son analizadas por un patólogo, quien realiza una evaluación cualitativa de la sobreexpresión mediante la inspección visual del tejido utilizando un microscopio, determina su clasificación según dos parámetros expresadas en variables lingüísticas: la intensidad de la coloración marrón (DAB), de débil a intensa, y la completitud, esto es, el porcentaje de la membrana que presenta tinción. A la muestra se asigna una clasificación, también referido como *scoring*, correspondiente a una de las siguientes cuatro categorías: 0, 1+ (negativos HER2), 2+ (equivocos) o 3+ (positivos HER2) [5]. Solo los pacientes con HER2 positivo pueden recibir el tratamiento dirigido a HER2.

Gracias al avance en los escáneres digitales de alta resolución, es posible digitalizar las muestras obteniendo una *whole-slide image* (WSI), imágenes de gran aumento y alta resolución, han permitido el desarrollo de *software* para el

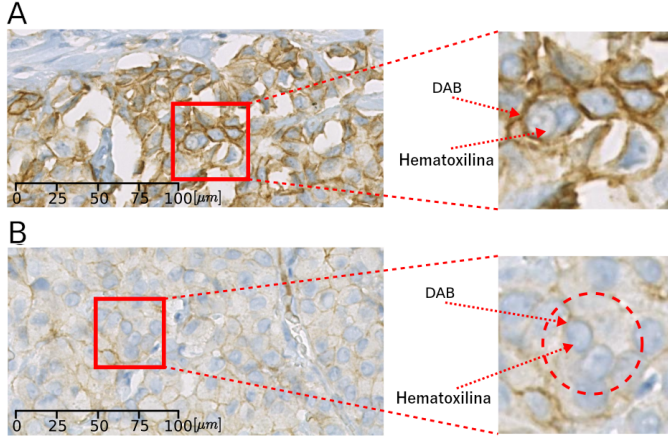


Fig. 1. Tejidos de dos pacientes tratados con IHC. La hematoxilina resalta los núcleos celulares con una tinción azul-violeta. DAB resalta membrana celular con una tinción marrón oscuro, cuando estas presentan sobreexpresión de HER2. El paciente (A) presenta un tejido con mayor sobreexpresión que el paciente (B), esto se aprecia en la intensidad del color marrón de las membranas celulares. En (A) se puede identificar fácilmente la cada célula y su membrana, mientras que en (B) esta tarea requiere un mayor esfuerzo. Sin importar el grado de la sobreexpresión, el núcleo celular presenta una tonalidad azul.

análisis computacional de estos tejidos. Sin embargo, desarrollar algoritmos para analizar estas imágenes es una tarea desafiante. Existen distintos factores que generaran variaciones, o ruido, en las distintas imágenes obtenidas. El proceso de tinción depende de la selección del anticuerpo y tiempo de fijación, estos provocan variaciones de color de las estructuras de interés o tinción no deseada. De igual forma, es desafiante la morfología irregular de la célula, la discontinuidad y las estructuras finas de las membranas, y la existencia de artefactos en la imagen.

Así, en este trabajo se abordará el problema de la segmentación semántica de imágenes y su aplicación práctica en escenarios médicos: segmentación semántica de estructuras celulares en imágenes de tejidos tratadas con IHC. Se utilizará el método de aprendizaje supervisado *decision tree* en su versión *fuzzy*. Esto permite la inclusión conocimiento del dominio, como la intensidad de la tinción DAB, en la construcción de variables.

Además, este problema de segmentación es un problema de clasificación no balanceado, por lo que requiere de un método con mayor sensibilidad a la clase relevante.

El resto de este documento está constituido por una descripción del método usado, *Fuzzy Decision Trees*, y su diferencia con el algoritmo clásico de *Decision Trees* en la sección II. La sección III detalle de la construcción del sistema de inferencia utilizado. La sección IV se resumen los resultados obtenidos. Finalmente, en la sección V termina con las conclusiones de este trabajo.

La implementación del *Fuzzy Decision Tree* y experimentos realizados se encuentran en el siguiente repositorio: github.com/sborquez/HER2_Fuzzy_Logic.

II. FUZZY DECISION TREES

A. Decision Tree

Los árboles de clasificación o *Decision Tree* (DT) son métodos de aprendizaje automático para construir modelos de predicción a partir de datos. Observan cómo las características individuales afectan el valor objetivo. Los modelos se obtienen particionando (subconjunto) recursivamente el espacio de datos y ajustar un modelo de predicción simple dentro de cada partición. Como resultado, la partición se puede representar gráficamente como un árbol de decisión. Los árboles de clasificación están diseñados para variables de destino que toman un finito número de valores desordenados, con el error de predicción medido en términos del costo de clasificación errónea.

B. Random Forest

Los árboles aleatorios o *Random Forest* (RF) es un algoritmo de clasificación, construido a partir de un conjunto de árboles de decisiones. Construye cada DT utilizando una muestra y subconjunto de características para crear un bosque de árboles no correlacionados. La predicción se construye a partir del consenso de la predicción de cada árbol.

C. Fuzzy Decision Tree with ID3 Algorithm

La variación fuzzy presentada en [6], denominado *Fuzzy Decision Tree* (FDT), permite incluir conocimiento de dominio al utilizar variables *fuzzy* como las variables de entrada del árbol. Junto a esto, las divisiones ya estaría predefinidas al momento de definir los *fuzzy sets*.

De esta forma, el algoritmo para ajustar los parámetros el árbol a los datos y el procedimiento para la predicción de nuevas instancias, varían en cierto grado al modelo clásico.

La versión presentada a continuación, se le agrega la posibilidad de considerar el desbalance de clases, es decir, cuando la razón del número de cada clases no es la misma.

1) *Notación:* Sea el T un FDT. T esta compuesto por nodos t . Un nodo que no tiene hijos se denomina hoja.

Sea el conjunto de datos D , donde cada dato x posee L atributos numéricos A_i , con $A_i \in \{A_1, \dots, A_L\}$ Cada dato x en D tiene asociado una etiqueta y , con el valor igual a una de las N clases distintas C_k , con $C_k \in \{C_1, \dots, C_N\}$. Para cada atributo A_i se definen M *fuzzy sets* (M variable para cada atributo A_i) y junto a sus funciones de membresía, F_{ij} y $\mu_{ij}(x)$ respectivamente. Se denomina $\mu^t(x)$ como el valor de membresía a un nodo t para un dato x . El valor será definido por el algoritmo.

Sea el subconjunto D^{C_k} , definido para cada nodo t , el conjunto de datos que cumple con un valor de membresía del nodo t mayor a 0 y pertenece a la clase C_k .

$$D^{C_k} : \{x | \mu^t(x) > 0 \wedge y = C_k\}$$

Sea el subconjunto $D_{F_{ij}}$, definido para cada nodo t , el conjunto de datos que cumple con una valor de membresía del *fuzzy set* F_{ij} mayor a 0.

$$D_{F_{ij}} : \{x | \mu_{F_{ij}}(x) > 0\}$$

Se introduce el vector pesos $w : \{w_{C_1}, \dots, w_{C_k}, \dots, w_{C_N}\}$, para el balanceo de importancia de clases, definido como:

$$w_{C_k} := \frac{1}{N} \frac{1}{\rho_k}$$

Donde ρ_k es la razón original de la clase C_k en el *dataset*.

Sea $|D^{C_k}|_w$ la cardinalidad del conjunto D^{C_k} , como la suma ponderada de los valores de membresía de los datos perteneciente al subconjunto D^{C_k} .

$$|D^{C_k}|_w = \sum_{x \in D^{C_k}} w_{C_k} \mu^t(x)$$

Utilizando el término anterior se define $|D|_w$.

$$|D|_w = \sum_{k \in [1, N]} |D^{C_k}|_w$$

De forma similar, se definen las cardinalidades respecto a los *fuzzy sets* de los atributos como sumas no ponderadas:

$$|D_{F_{ij}}| = \sum_{x \in D_{F_{ij}}} \mu_{F_{ij}}(x)$$

$$|D| = \sum_{j \in [1, M]} |D_{F_{ij}}|$$

2) *Construcción del Árbol*: Para construir un FDT, se debe seguir los siguientes pasos:

Dado un conjunto de datos D , el vector de pesos w , los parámetros θ_n y θ_r , y los *fuzzy set* y sus funciones de membresía F_{ij} y μ_{ij}

Paso 1: Se genera una raíz t con el conjunto de datos D , se definen los valores de membresía $\mu^t(x)$ igual a 1 para cada valor x en D .

Paso 2: Si el nodo t cumple con una de las siguientes tres condiciones, se considera una hoja.

- No quedan más atributos para comparar.
- La cantidad de datos es menor a θ_n :

$$|D| \leq \theta_n$$

- Si para alguna clase C_k se cumple:

$$\frac{|D^{C_k}|_w}{|D|_w} > \theta_r$$

Paso 3: En caso contrario, para cada atributo A_i se obtiene la ganancia $G(A_i, D)$, y se elige el atributo con mayor ganancia, $A_{\text{máx}}$. Donde:

$$G(A_i, D) = I(D) - E(A_i, D)$$

$$I(D) = - \sum_{k \in [1, N]} (p_k \log_2(p_k))$$

$$E(A_i, D) = \sum_{j \in [1, M]} (p_{ij} \cdot I(D_{F_{ij}}))$$

$$p_k = \frac{|D^{C_k}|_w}{|D|_w}$$

$$p_{ij} = \frac{|D_{F_{ij}}|}{|D|}$$

Paso 4: Se subdivide D en M subconjunto de acuerdo al atributo $A_{\text{máx}}$, donde cada nuevo D_j se corresponde con uno de los *fuzzy set* $F_{\text{máx},j}$, cada x en D_j debe cumplir con $\mu_{\text{máx},j} > 0$.

Paso 5: Se generan M nuevos nodos t_j para cada D_j . Con nuevos valores de membresía $\mu^{t_j}(x) = \mu^t(x) \cdot \mu_{\text{máx},j}(x)$. Luego, se repite el **paso 2** para cada t_j .

Al definir un nodo como hoja, se obtienen los valores p_k , estos representan la membresía de cada clase C_k .

3) *Razonamiento de clase con FDT*: A diferencia del algoritmo clásico de DT, un FDT no solo visita un camino desde la raíz a un hoja, en cambio, obtiene los valores de membresía para cada *fuzzy set* de un atributo en un nodo, y los propaga a los nodos siguientes, multiplicando cada valor hasta llegar a cada una de las hojas.

Luego, se puede obtener el valor de membresía para una clase C_k , como la suma ponderada de p_k de cada hoja, ponderada por la membresía de cada hoja.

La figura 2 muestra este procedimiento para un árbol con dos clases C_1 y C_2 .

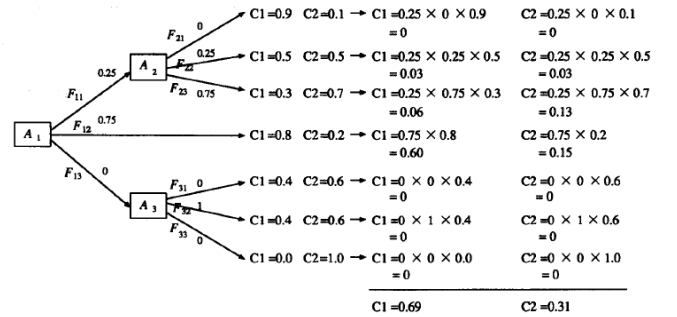


Fig. 2. Razonamiento con *Fuzzy Decision Tree*.

III. SISTEMA DE INFERENCIA FUZZY PARA SEGMENTACIÓN DE IMÁGENES

Para cumplir con la tarea de segmentación se plantea el siguiente sistema o *Pipe*, compuesto por los siguientes componentes:

- Extracción de características.
- Fuzzificación

- Fuzzy Decision Tree
- Post-Procesado

A. Extracción de Características

Este componente corresponde a la aplicación de transformación, filtros y operaciones sobre la imagen de entrada para obtener una serie de descriptores a nivel de píxel. A continuación se lista cada atributo extraído.

- Atributos de Color. Valores promedio de un vecindario 3x3.
 - Intensidades RGB.
 - Intensidades HSV.
 - Intensidad Escala de Grises.
 - Exceso RGB (i.e. $2R - (G+B)$)
 - Deconvolución Haematoxylin-Eosin-DAB [7] (HED)
- Atributos de Filtros [8]:
 - Filtro Laplaciano.
 - Magnitud gradiente de Sobel.
- Atributos de Textura.
 - GLCM (Gray Level Co-occurrences-Matrix) [9]
 - Local Binary Patterns, rotation invariant extension. [10]

Para la selección final de atributos, se utilizó un gráfico de importancia de características de un *Random Forest* ajustado utilizando todos los atributos (ver figura 3)

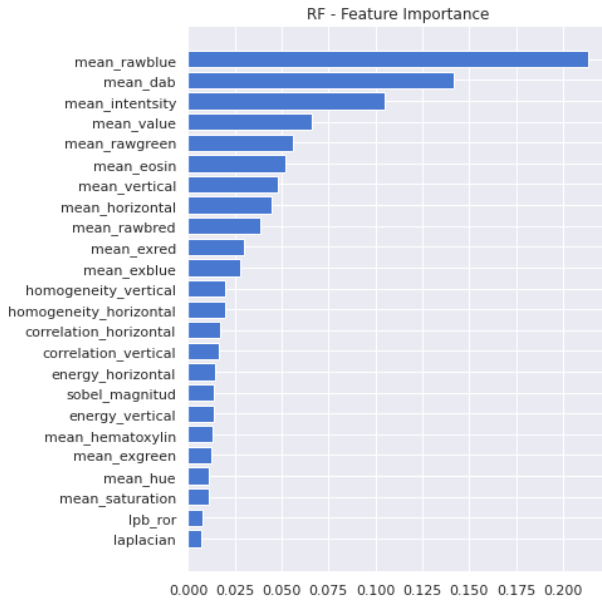


Fig. 3. Importancia de atributos de un *Random Forest*.

Finalmente, se utilizan dos *sets* de atributos:

- 1) Set 1: Definido por RF.
 - a) mean_rawblue
 - b) mean_dab
 - c) mean_intensity
 - d) mean_rawgreen
 - e) mean_eosin

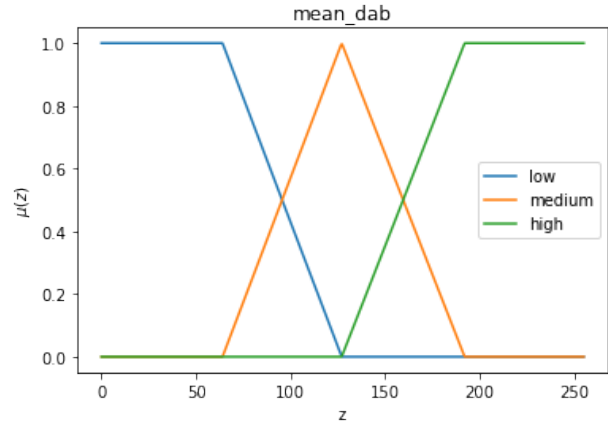


Fig. 4. Fuzzy sets del atributo DAB.

- f) mean_vertical
- g) mean_rawred
- h) homogeneity_vertical
- i) mean_hematoxylin
- j) sobel_magnitud

- 2) Set 2: Definido por atributos relevantes para el experto [5].

- a) mean_rawblue
- b) mean_dab
- c) mean_intensity
- d) homogeneity_vertical
- e) sobel_magnitud

B. Fuzzificación de Características

Para este trabajo solo se consideraron funciones sigma, triangulares y trapezoidales [11]. Sin embargo, esta selección de *fuzzy sets* se deja en libertad y conocimiento del experto. La figura 4 muestra un ejemplo de estas funciones, para el atributo *mean_dab*.

C. Segmentación usando FDT

El problema de segmentación puede entenderse como un problema de clasificación binaria, donde se busca clasificar cada píxel en una de las dos clases: *overexpressed* (clase positiva) y *non-overexpressed* (clase negativa). Esto nos permite utilizar la construcción de FDT visto en la sección II, utilizando los atributos y *fuzzy sets* obtenidos en los componentes anteriores como entrada del FDT.

Sin embargo, este es un problema de clases desbalanceadas, como se puede observar en la figura 5, hay una clara mayoría de instancias de la clase negativa *non-overexpressed*.

Con el objetivo de determinar la importancia de los atributos y desbalance de clases, se entrenan tres FDT:

- FDT A** Atributos *Set 1*, no considera desbalance.
- FDT B** Atributos *Set 1*, si considera desbalance.
- FDT C** Atributos *Set 2*, si considera desbalance.

Los tres árboles comparten los siguientes parámetros: $\max_deep = 4$, $\theta_r = 0.98$ y $\theta_n = 15$.

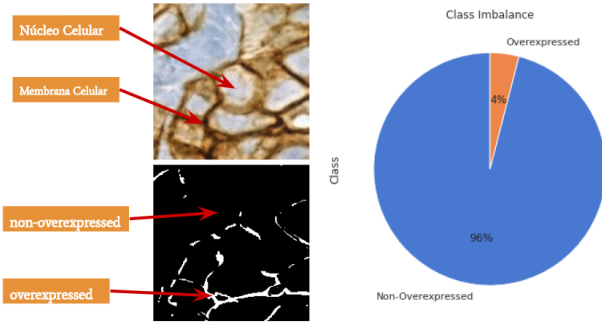


Fig. 5. Segmentación de clases, de color blanco la clase positiva *overexpressed*, de color negro la clase negativa *non-overexpressed*. El gráfico de la derecha muestra el desbalance de clases.

TABLA I
DATASET UTILIZANDO.

Dataset	Train	Test
Imágenes	1+_2.tif	1+_25.tif 2+_9.tif 3+_15.tif
	1+_20.tif	
	2+_1.tif	
	2+_8.tif	
	3+_16.tif	
	3+_19.tif	
Píxeles	6.000.000	3.000.000

D. Post-Procesamiento

La construcción de la máscara de segmentación resultante, para cada salida del FDT, se elige la clase con mayor valor de membresía. Luego, se reestructura el vector de salida en una imagen bidimensional.

IV. EVALUACIÓN

El *dataset* de imágenes disponibles está conformado por imágenes de una resolución de 1000x1000 píxeles. Junto a cada imagen, se tiene su máscara de segmentación, la etiqueta para cada píxel. Las imágenes se dividen en los conjuntos de *train* y *test*, como indica la tabla I

La métricas utilizadas corresponden a las medidas de *precision*, *recall*, *F1-Score* y *Dice* definidas en el anexo A. Dado que es un problema de clasificación binaria con clases desbalanceadas, se dará mayor importancia a la métrica de *recall*, ya que es un indicador que una buena clasificación de la clase relevante (*overexpressed*).

TABLA II
RESULTADOS DE ENTRENAMIENTOS DE ÁRBOLES FDT A Y FDT B.

Métrica	Precision	Recall	F1-Score	Dice	Soporte
FDT A (no considera desbalance)					
overexpressed	0,77	0,25	0,38	0,381	211.306
non-overexpressed	0,95	0,99	0,97	-	2.788.694
FDT B (sí considera desbalance)					
overexpressed	0,37	0,93	0,53	0,531	211.306
non-overexpressed	0,99	0,88	0,93	-	2.788.694

TABLA III
RESULTADOS DE ENTRENAMIENTO DE MÉTODOS FUZZY Y CRISP.

Métrica	Precision	Recall	F1-Score	Dice	Soporte
FDT B (Set 1)					
overexpressed	0,37	0,93	0,53	0,521	211.306
non-overexpressed	0,99	0,88	0,93	-	2.788.694
FDT C (Set 2)					
overexpressed	0,35	0,95	0,52	0,517	211.306
non-overexpressed	1,00	0,87	0,93	-	2.788.694
Decision Tree (Set 1)					
overexpressed	0,97	0,04	0,08	0,08	211.306
non-overexpressed	0,93	1,00	0,96	-	2.788.694
Random Forest (Set 1)					
overexpressed	0,46	0,87	0,60	0,601	211.306
non-overexpressed	0,99	0,92	0,95	-	2.788.694

A. Balanceo de clases usando el vector w

A continuación, se analiza el impacto de usar el vector de pesos w en un problema de clases no balanceadas. Para esta sección, se comparan los resultados de **FDT A** y **FDT B**.

Los resultados están en la tabla II. Los resultados indican que la versión balanceada, **FDT B**, obtiene mejores resultados que **FDT A**. Las máscaras obtenidas se encuentran en el A.

B. Fuzzy Decision Tree

Se compara el desempeño de **FDT B** y **FDT C** frente a los modelos **DT** y **RF**. En este caso, los cuatro modelos consideran el desbalance de clases.

Los resultados se encuentran en la tabla III, los resultados demuestran una mejora en respecto al *Recall*. Sin embargo, el uso de los modelos fuzzy, requieren mucho más tiempo para realizar inferencia para cada píxel. Esto se debe a que FDT visitar todas las hojas del árbol para realizar la inferencia.

Por otro lado, el los atributos utilizados llegan a resultados similares, es más, los árboles tienden a utilizar las mismas reglas de inferencia.

V. CONCLUSIONES Y COMENTARIOS

La modificación del método de FDT con clases desbalanceadas demostró tener un gran impacto en la sensibilidad de la clase positiva, obteniendo un *recall* 2.5x veces que el FDT que no considera el desbalance.

Los resultados obtenidos demuestran la competencia del uso de *Fuzzy Decision Tree* en el ámbito de la segmentación de imágenes. Los métodos de *fuzzy* obtienen un desempeño comparable con los métodos *crisp*. Estos tienen mejores resultados en *recall* para la clase positiva, sin embargo, no logran superar el desempeño en las métricas *f1-score* y *dice* obtenido por el RF.

Una segunda observación se puede obtener al analizar el tiempo de predicción que requiere cada modelo para una entrada, se puede notar que el tiempo requerido por los modelos FDT, en comparación con el modelo DT, es 2200x más lento.

Un segundo resultado que se puede derivar de los caminos desde la raíz del FDT hasta una de sus hojas. Este camino permite la construcción de una regla de inferencia *fuzzy*, al conjugar cada cada nodo y *fuzzy set* involucrado en el camino.

Por ejemplo:

IF ("mean_rawgreen" IS "medium") AND ("mean_rawblue" IS "medium") AND ("mean_dab" IS "high"), THEN IS "overexpressed"

Al realizar esta construcción para cada camino que alcanza una hoja, se obtiene un conjunto de reglas de inferencia.

Como trabajo futuro, es posible realizar un análisis de este conjunto de reglas derivado, por ejemplo, determinar el subconjunto mínimo de reglas que permitan realizar la clasificación, si perder una sensibilidad o desempeño relevante.

REFERENCIAS

- [1] C. Gutierrez and R. Schiff, "Her2: biology, detection, and clinical implications," *Archives of pathology & laboratory medicine*, vol. 135, no. 1, pp. 55–62, 2011.
- [2] M. M. Moasser, "The oncogene her2: its signaling and transforming functions and its role in human cancer pathogenesis," *Oncogene*, vol. 26, no. 45, pp. 6469–6487, 2007.
- [3] J. Wang and B. Xu, "Targeted therapeutic options and future perspectives for her2-positive breast cancer," *Signal transduction and targeted therapy*, vol. 4, no. 1, pp. 1–22, 2019.
- [4] R. Pezoa, "A computational method for segmentation of histopathological images with her2 overexpression," *Universidad Técnica Federico Santa María*, vol. 1, pp. 5–17, 2017.
- [5] A. C. Wolff, M. E. H. Hammond, K. H. Allison, B. E. Harvey, P. B. Mangu, J. M. Bartlett, M. Bilous, I. O. Ellis, P. Fitzgibbons, W. Hanna, R. B. Jenkins, M. F. Press, P. A. Spears, G. H. Vance, G. Viale, L. M. McShane, and M. Dowsett, "Human epidermal growth factor receptor 2 testing in breast cancer: American society of clinical oncology/college of american pathologists clinical practice guideline focused update," *Journal of Clinical Oncology*, vol. 36, no. 20, pp. 2105–2122, 2018, pMID: 29846122. [Online]. Available: <https://doi.org/10.1200/JCO.2018.77.8738>
- [6] M. Umanol, H. Okamoto, I. Hatono, H. Tamura, F. Kawachi, S. Umedzu, and J. Kinoshita, "Fuzzy decision trees by fuzzy id3 algorithm and its application to diagnosis systems," in *Proceedings of 1994 IEEE 3rd International Fuzzy Systems Conference*. IEEE, 1994, pp. 2113–2118.
- [7] A. C. Ruifrok, D. A. Johnston *et al.*, "Quantification of histochemical staining by color deconvolution," *Analytical and quantitative cytology and histology*, vol. 23, no. 4, pp. 291–299, 2001.
- [8] R. C. Gonzalez and E. Richard, "Woods, digital image processing," *ed: Prentice Hall Press, ISBN 0-201-18075*, vol. 8, 2002.
- [9] M. Hall-Beyer, "Glc texture: a tutorial," *National Council on Geographic Information and Analysis Remote Sensing Core Curriculum*, vol. 3, 2000.
- [10] T. Ojala, M. Pietikainen, and T. Maenpaa, "Multiresolution gray-scale and rotation invariant texture classification with local binary patterns," *IEEE Transactions on pattern analysis and machine intelligence*, vol. 24, no. 7, pp. 971–987, 2002.
- [11] K. Tanaka, "An introduction to fuzzy logic for practical applications," 1997.

APPENDIX

Métricas de Clasificación:

Sea TP, TN, FP y FN las abreviaciones para *true positive*, *true negative*, *false positive* y *false negative* respectivamente.

Se definen las siguientes métricas de clasificación:

$$\begin{aligned} \text{precision} &= \frac{TP}{TP + FP} \\ \text{recall} &= \frac{TP}{TP + FN} \\ \text{f1-score} &= 2 * \frac{\text{precision} \cdot \text{recall}}{\text{precision} + \text{recall}} \end{aligned}$$

Métrica de Segmentación Binaria:

A diferencia de las métricas de clasificación, el cálculo se obtiene de las máscaras de segmentación generada.

Sean P y T las máscara de segmentación binaria obtenida por un modelo y la anotada por el experto, respectivamente. Se utiliza el operador \cap como la intersección de las máscaras, es decir, el conjunto de píxeles clasificados positivos en ambas máscaras.

$$\text{DICE} = \frac{2 * |P \cap T|}{|P| + |T|}$$

Resultados en Test Set

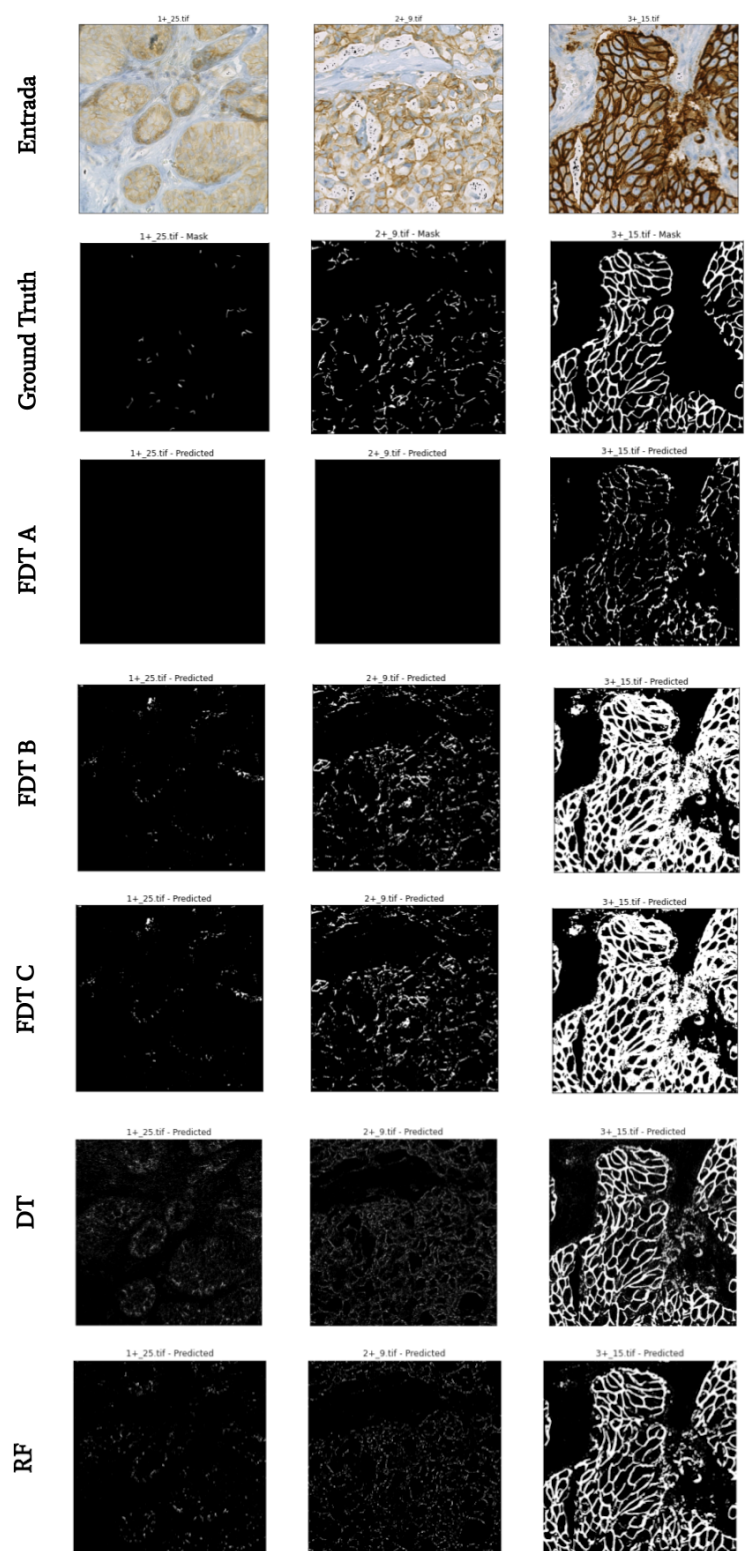


Fig. 6. Resultados de métodos de segmentación en el conjunto Test.