

The background of the slide is a histopathological image, likely a micrograph of tissue. It shows a complex network of dark, branching structures (possibly blood vessels or ducts) set against a lighter, more granular background. The overall color palette is dominated by dark brown and black tones, with some lighter, almost white, areas. The texture is highly detailed and organic.

INF-575 Fuzzy Logic

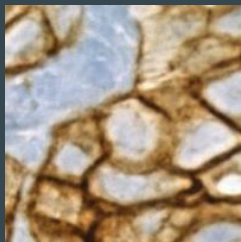
Segmentación de Imágenes Histopatológicas con Fuzzy Decision Trees

Sebastián Bórquez G.

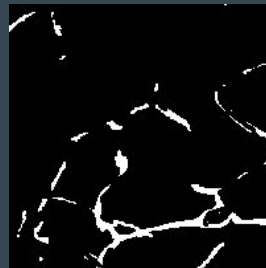
Objetivos y Metodología

Objetivo

Implementar un sistema fuzzy para la segmentación de la membrana celular en imágenes de tejidos tratados con Inmunohistoquímica.



Inferencia Fuzzy para
la segmentación de
imágenes



Metodología

Comparación del desempeño de 4 modelos en la tarea de clasificación:

- Decision Tree Classifier
- Random Forest
- Fuzzy Decision Tree (class weighted)
- Fuzzy Decision Tree (without class weight)

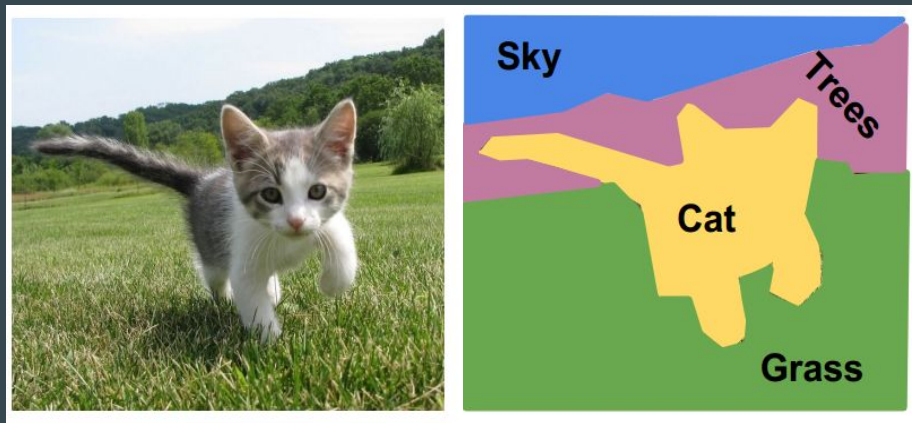
Utilizando métricas de clasificación e inspección visual.

Antecedentes

Segmentación de Imágenes

La segmentación de imágenes corresponde a la división de una imagen digital en varias partes u objetos.

Puede entender como un **problema de clasificación** a nivel de píxel de la imagen.



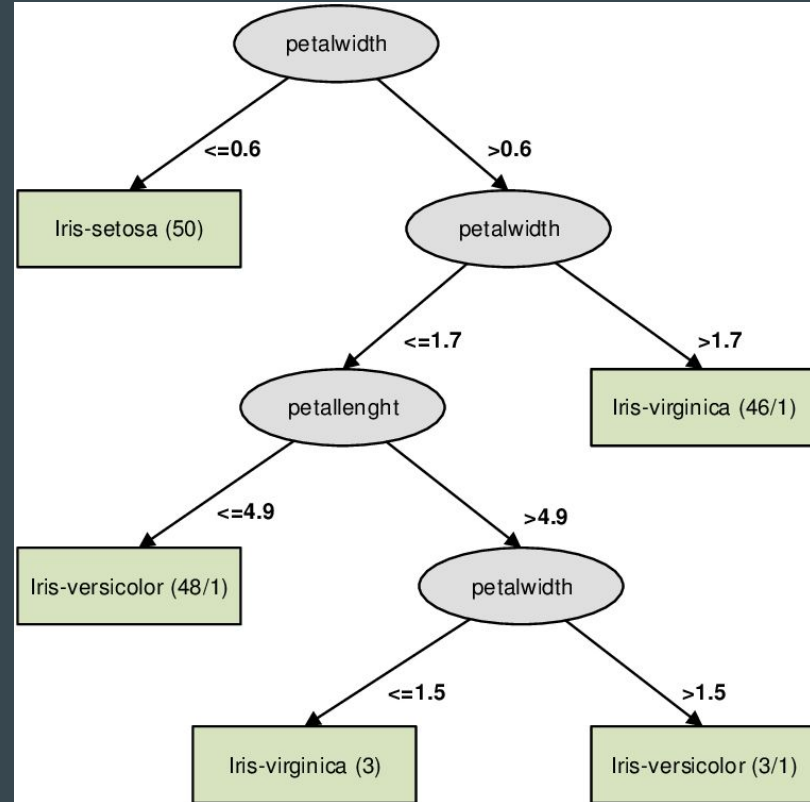
Árboles de Decisión

Un árbol de decisión es un modelo de clasificación.

Compuesto por nodos. Utiliza uno de los atributos para determina su pertenencia en uno de los conjuntos (crisp) existentes.

Si el nodo es una hoja, esta determinará la clase a la que pertenece.

Entrenar un árbol es determinar qué atributos por nodo y divisiones (conjuntos) óptimas.



Fuzzy Decision Trees by Fuzzy ID3 Algorithm

Un árbol de decisión fuzzy se diferencia en lo siguiente:

- Las “divisiones” de cada atributos están pre-definidas por conjuntos fuzzy.
- Al entrenar, se determina los atributos A_i que maximicen la ganancia de información G .
- Durante la predicción, la selección de conjuntos (o ramas) realizada en cada nodo no es única, ya que utiliza el grado de pertenencia para obtener la clase.
- Permite la construcción de reglas.

Fuzzy Decision Trees by Fuzzy ID3 Algorithm

Definición: Un árbol está compuesto por nodos y hojas

Sea el conjunto de datos D .

Donde cada dato tiene L atributos numéricos $A_i: \{A_{i1} \dots A_{iL}\}$

Y asignado una clase C_k . De un conjunto de N clases.

Cada atributo tiene M fuzzy sets (variable de cada atributo), F_{im} y sus funciones de membresía μ_{im} .

Sea D^{C_k} el subconjunto de datos de la clase. $D_{F_{ij}}$ el subconjunto de datos pertenecientes al fuzzy set F_{ij}

Sea $|D|_w$ y $|D^{C_k}|_w$ la suma ponderada con w de las funciones de membresía de un conjunto de datos.

Fuzzy Decision Trees by Fuzzy ID3 Algorithm

Construcción:

1. Se genera una raíz con todos los datos D , y una membresía igual a 1 para cada dato.
2. Si el nodo actual t cumple con una de las tres condiciones, es considerado hoja.
 - a. Si para alguna clase C_k , la proporción de clases es mayor o igual a θ_r .

$$(|D^{C_k}|_w) / |D|_w \geq \theta_r$$

- b. La cantidad de miembros es menor a θ_n .

$$|D|_w < \theta_n$$

- c. No quedan más atributos.

Fuzzy Decision Trees by Fuzzy ID3 Algorithm

Construcción:

3. En caso contrario, se calcula la ganancia para cada atributo A_i

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

Donde:

$$I(D) = -\sum_{k=1}^N (p_k \log_2 p_k)$$

$$E(A_i, D) = \sum_{k=1}^M (p_k I(D_{Fij}))$$

$$p_k = |D^{ck}| / |D|$$

$$p_{ij} = |D_{Fij}| / \sum_{j=1}^M |D_{Fij}|$$

A_{\max} es el atributo con mayor ganancia.

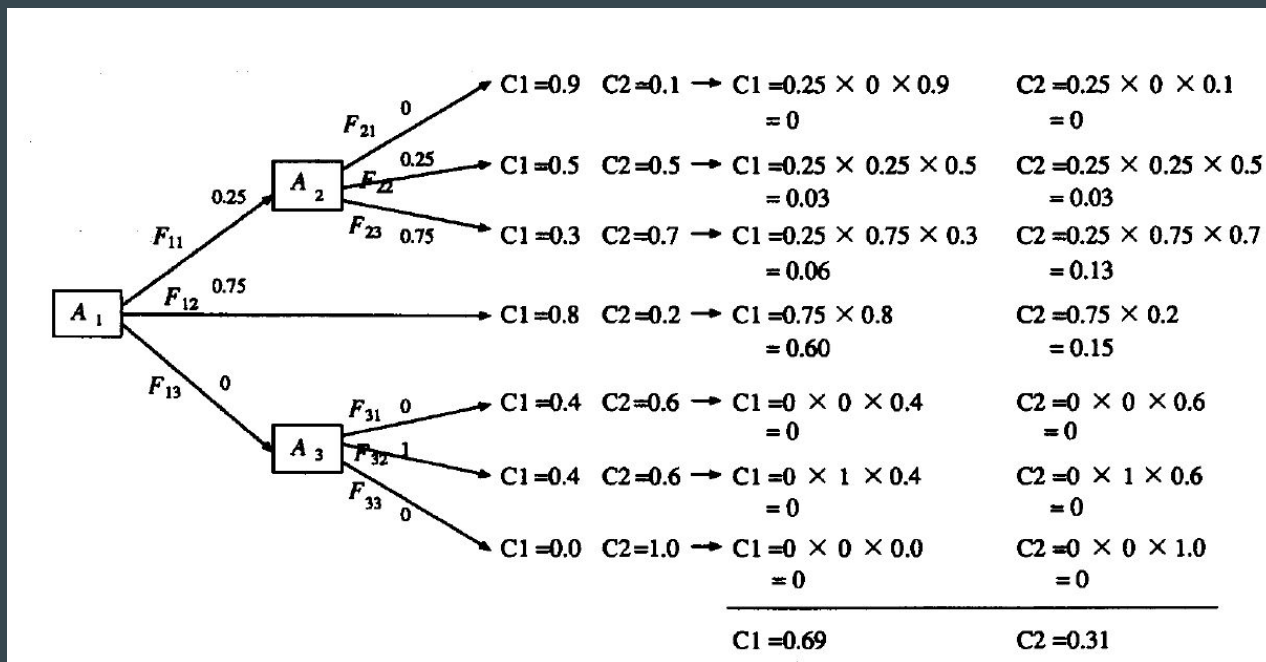
Fuzzy Decision Trees by Fuzzy ID3 Algorithm

Construcción:

4. Se subdivide D en M subconjuntos $D_1, \dots, D_j, \dots, D_M$ de acuerdo al atributo A_{\max} . Y su nuevo valor de membresía $\mu_{\max,j}$ corresponde al producto del valor de membresía en D y de $F_{\max,j}$.
5. Generar nuevos nodos t_j para cada D_j y repetir desde 2.

Fuzzy Decision Trees by Fuzzy ID3 Algorithm

Predicción



Fuzzy Decision Trees by Fuzzy ID3 Algorithm

iris setosa



petal

sepal

iris versicolor



petal

sepal

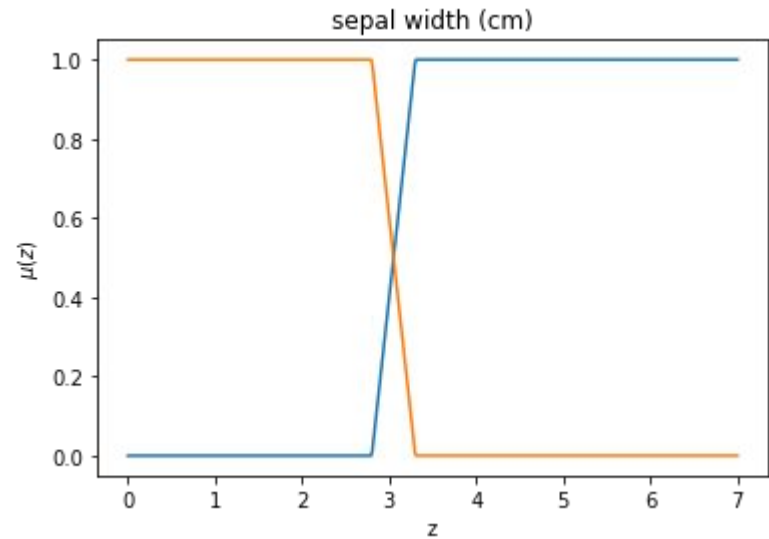
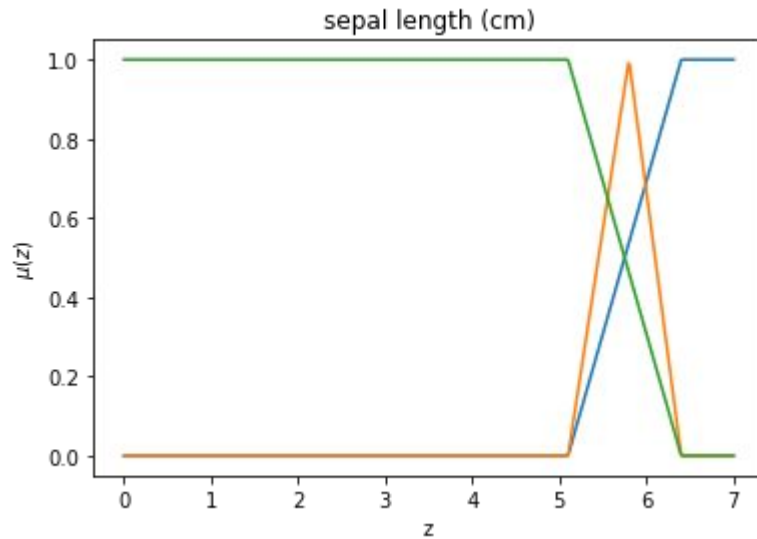
iris virginica



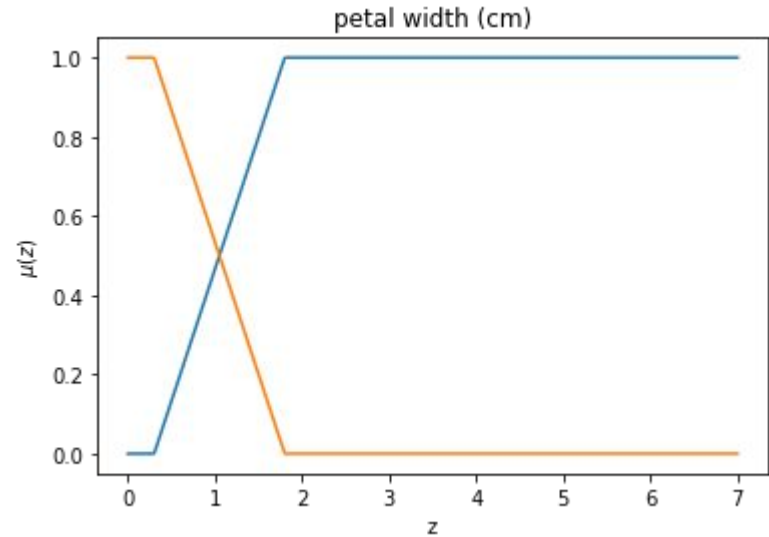
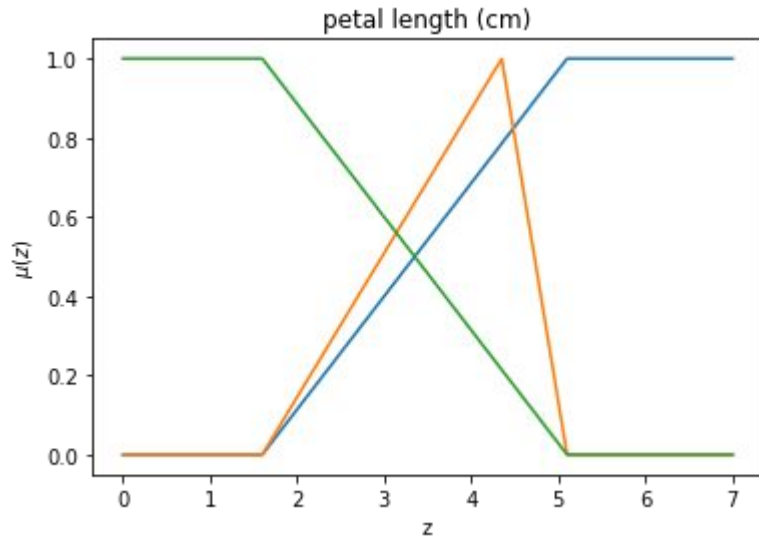
petal

sepal

Fuzzy Decision Trees by Fuzzy ID3 Algorithm



Fuzzy Decision Trees by Fuzzy ID3 Algorithm



Fuzzy Decision Trees by Fuzzy ID3 Algorithm

```
petal length (cm):  
  [high]-> petal width (cm):  
    [high]-> sepal length (cm):  
      [high]-> sepal width (cm):  
        [high]-> {'setosa': 0.0, 'versicolor': 0.24, 'virginica': 0.76}  
        [low]-> {'setosa': 0.0, 'versicolor': 0.3, 'virginica': 0.7}  
      [medium]-> sepal width (cm):  
        [high]-> {'setosa': 0.0, 'versicolor': 0.61, 'virginica': 0.39}  
        [low]-> {'setosa': 0.0, 'versicolor': 0.49, 'virginica': 0.51}  
      [low]-> sepal width (cm):  
        [high]-> {'setosa': 0.01, 'versicolor': 0.67, 'virginica': 0.33}  
        [low]-> {'setosa': 0.0, 'versicolor': 0.58, 'virginica': 0.42}  
    [low]-> {'setosa': 0.02, 'versicolor': 0.9, 'virginica': 0.08}  
  [medium]-> {'setosa': 0.01, 'versicolor': 0.91, 'virginica': 0.08}  
  [low]-> petal width (cm):  
    [high]-> {'setosa': 0.09, 'versicolor': 0.84, 'virginica': 0.07}  
    [low]-> {'setosa': 0.91, 'versicolor': 0.09, 'virginica': 0.0}
```

Fuzzy Decision Trees by Fuzzy ID3 Algorithm

Ejemplos:

IF (("petal length (cm)" IS "high") AND ("petal width (cm)" IS "high") AND ("sepal length (cm)" IS "high") AND ("sepal width (cm)" IS "low")), THEN IS "virginica"

IF (("petal length (cm)" IS "high") AND ("petal width (cm)" IS "high") AND ("sepal length (cm)" IS "low") AND ("sepal width (cm)" IS "high")), THEN IS "versicolor"

IF (("petal length (cm)" IS "low") AND ("petal width (cm)" IS "low")), THEN IS "setosa"

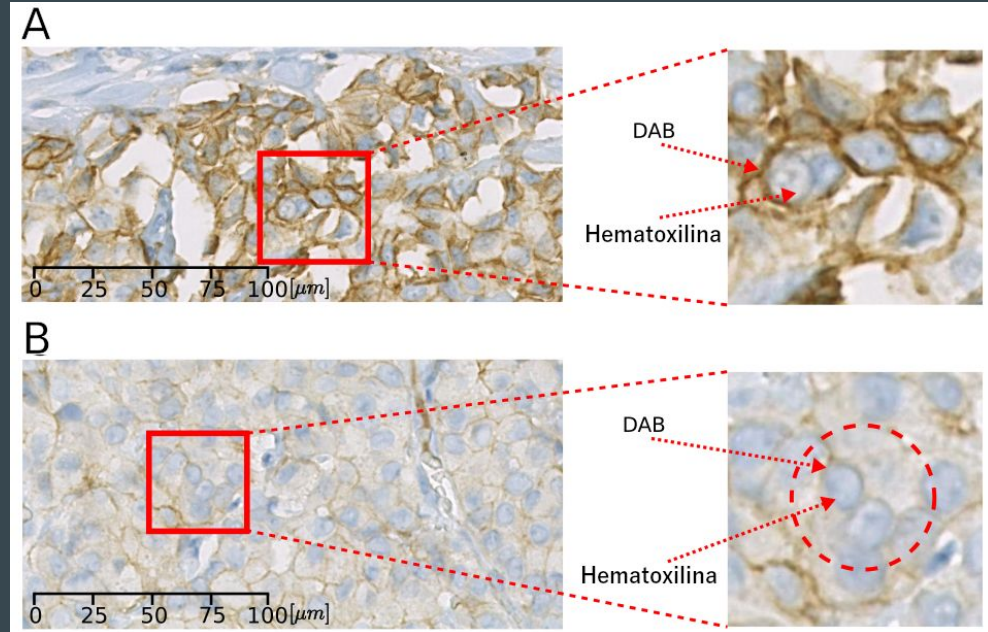
Contexto

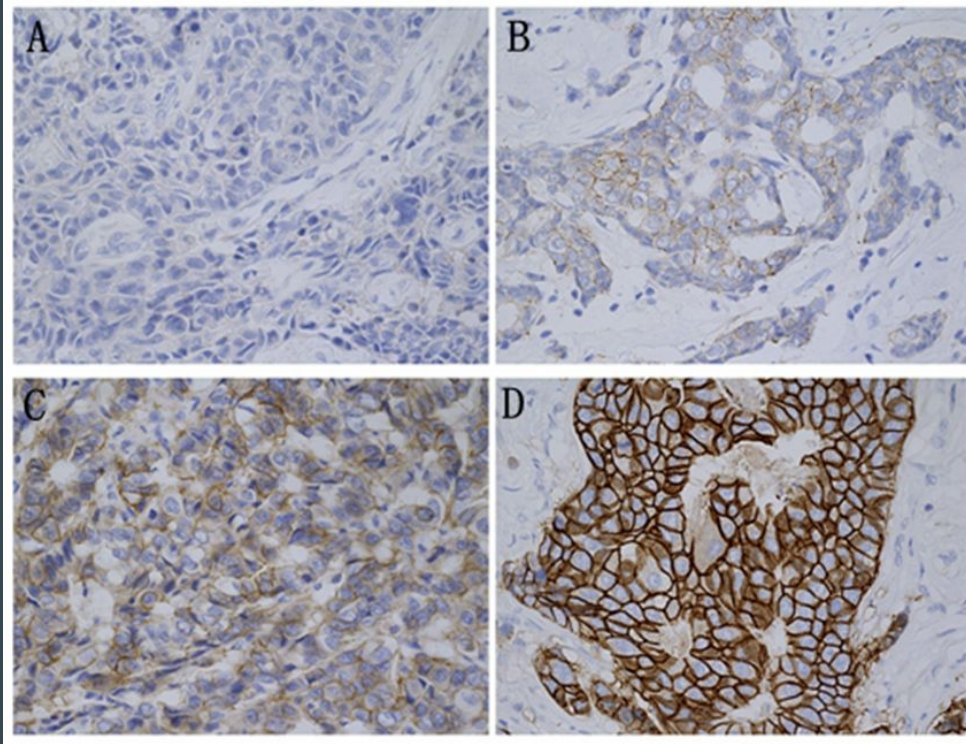
HER2 Scoring

La proteína HER2 (*human epidermal growth factor receptor 2*) está presente de manera natural en nuestras células.

Estudios han demostrado la correlación de su sobreexpresión con ciertos tipos de cáncer.

El análisis de imágenes de tejido permite determinar la presencia de sobreexpresión.

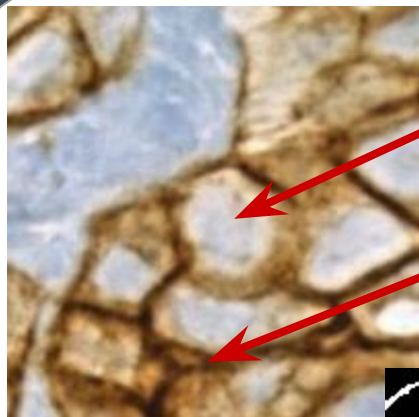




Según la guía de clasificación de la American Society of Clinical Oncology/ College of American Pathologists Clinical Practice Guideline Focused Update.

- A) 0+ Negativo
- B) 1+ Negativo
- C) 2+ Equivoco
- D) 3+ Positivo

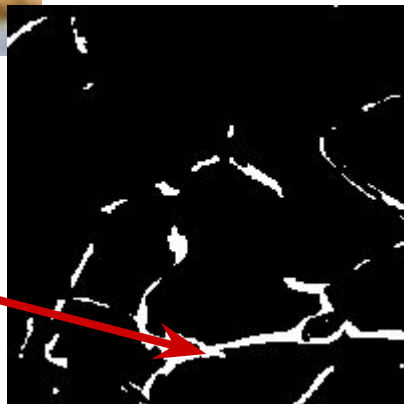
HER2 Scoring



Núcleo Celular

Membrana Celular

Sobreexpresión

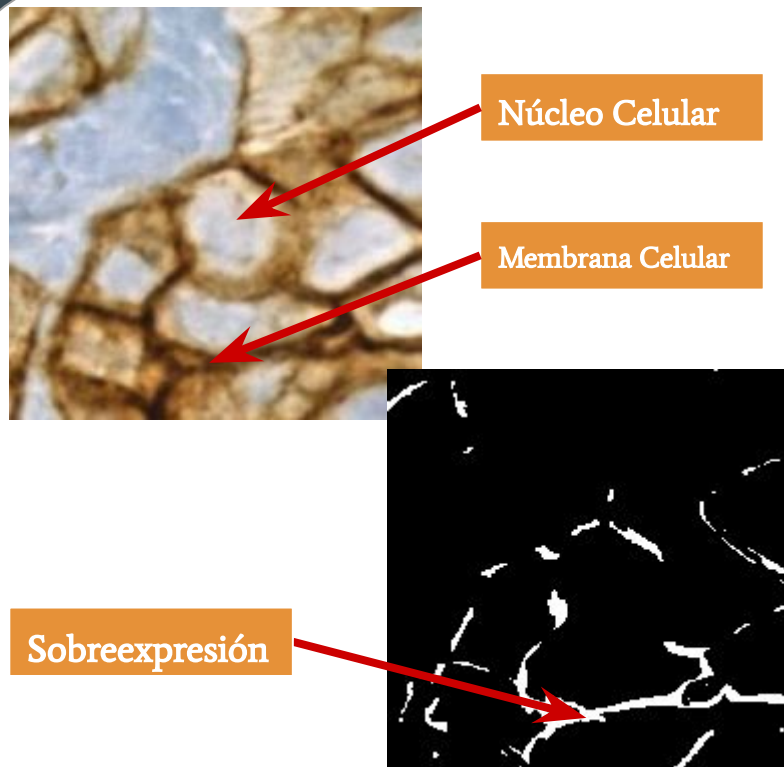


Tejidos tratados con inmunohistoquímica, permiten resaltar estructuras con sobreexpresión de la proteína.

La sobreexpresión de la proteína HER2 presenta tonalidades intensas en la membrana celular.

De esta forma, realizar **una segmentación de la membrana celular**, permite cuantificar la **sobreexpresión**.

HER2 Scoring

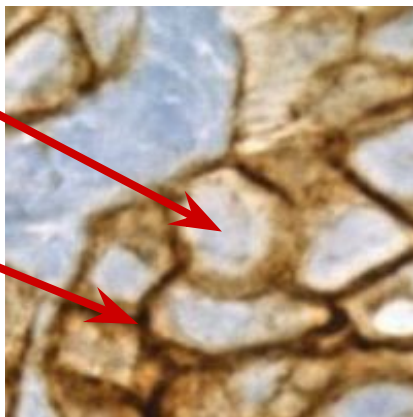


Clasificar cada píxel en una de las dos clases:

- **Sobreexpresión**
- **Sin-Sobreexpresión**

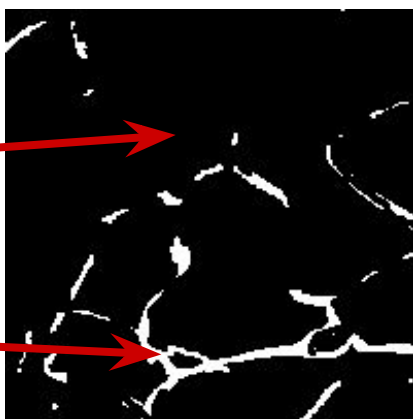
Núcleo Celular

Membrana Celular



non-overexpressed

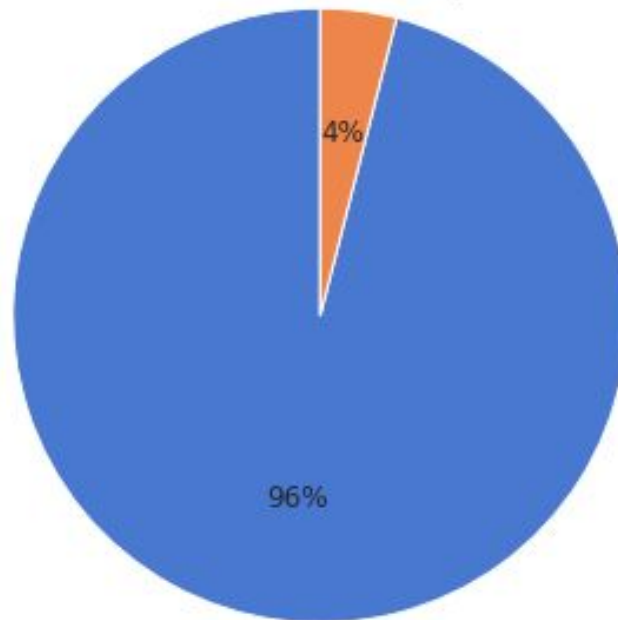
overexpressed



Class

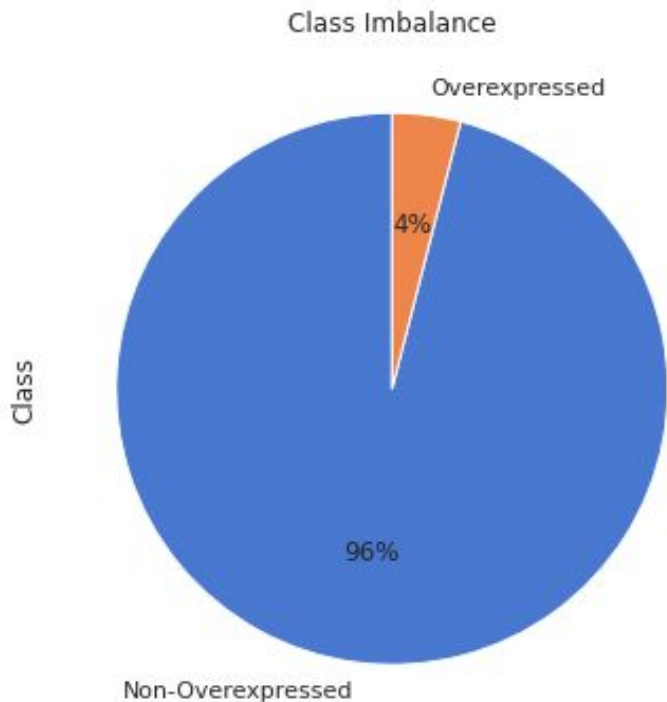
Class Imbalance

Overexpressed



Non-Overexpressed

HER2 Scoring



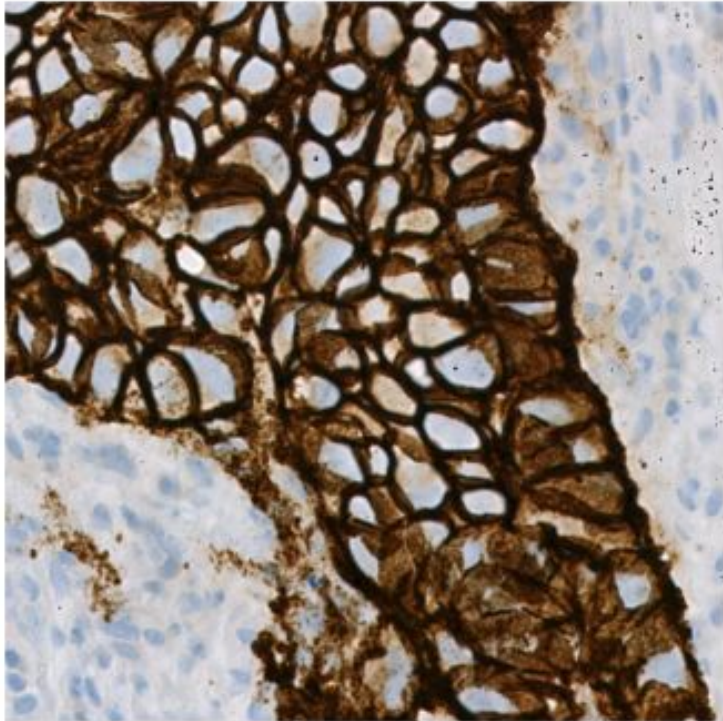
Clasificar cada píxel en una de las dos clases:

- **Sobreexpresión**
- **Sin-Sobreexpresión**

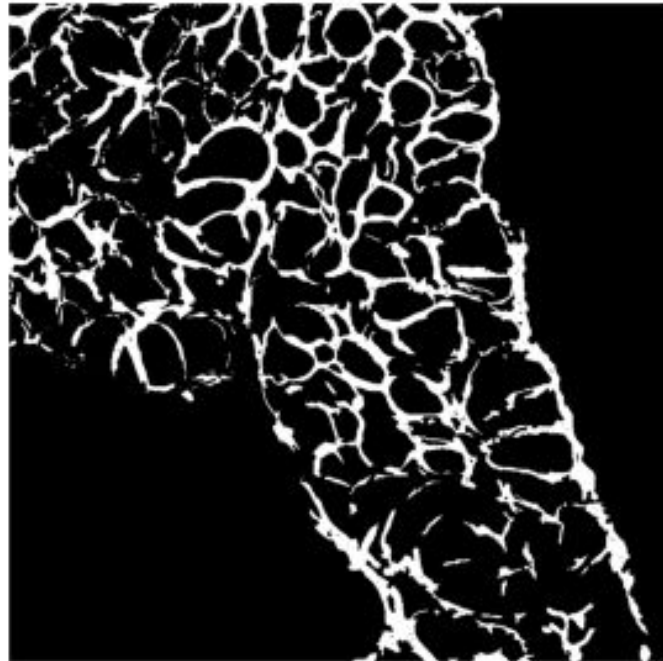
Clases no balanceadas

Recursos

3+_16.tif



3+_16.tif - Mask



Recursos

Train Images:

1+_2.tif,
1+_20.tif
2+_1.tif
2+_8.tif
3+_16.tif
3+_19.tif

Test Images:

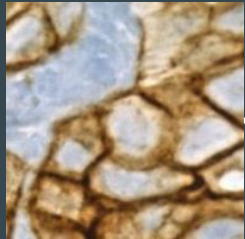
1+_25.tif,
2+_9.tif
3+_15.tif

- Resolución:
1000x1000
- Máscara con 2
clases :
0: Non-Overexpression
1: Overexpression

Solución Propuesta

Sistema de Inferencia Fuzzy

Podemos generar un sistema de inferencia basado en *Fuzzy Decision Trees* para realizar una segmentación de la imagen.



Extracción de
características

Fuzzificación

Fuzzy Decision
Tree

Post procesado



Extracción de Características

Extracción de
características

Obtener de la imagen, medidas y/o características que permitan abstraerse de la imagen.

Nos interesan características a nivel de píxel (o vecindario):

- Intensidades en diferentes espacios de colores.
- Filtros/Derivadas
- Textura.

El objetivo es tener nuevas variables para la fuzzificación.



Extracción de Características

Extracción de
características

Características de Color

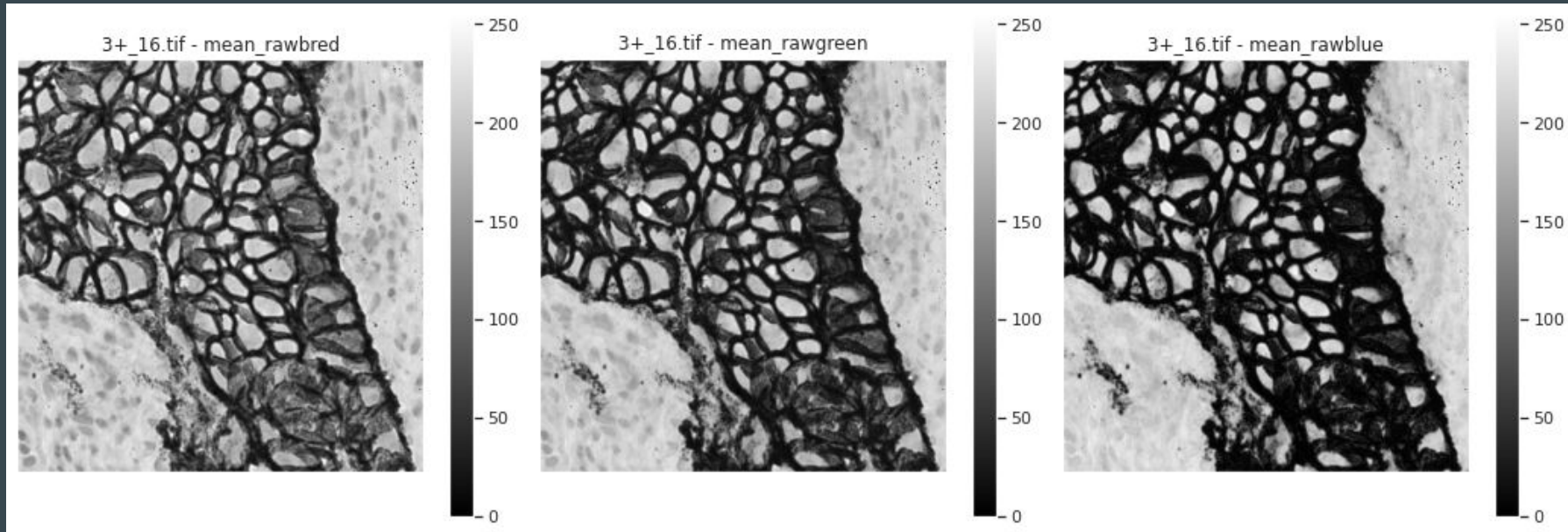
Se obtienen a partir del promedio de un pequeño vecindario de píxeles de tamaño 3x3:

- RGB
- Gray
- Haematoxylin-Eosin-DAB (HED)
- HSV
- RGB Excess
 - $2R - (G+B)$
 - $2G - (R+B)$
 - $2B - (G+R)$

Extracción de Características

Extracción de
características

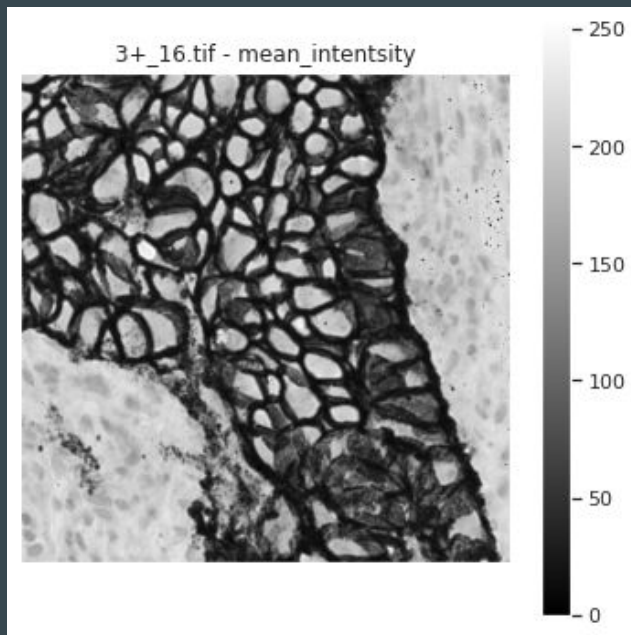
Características de Color: RGB



Extracción de Características

Extracción de
características

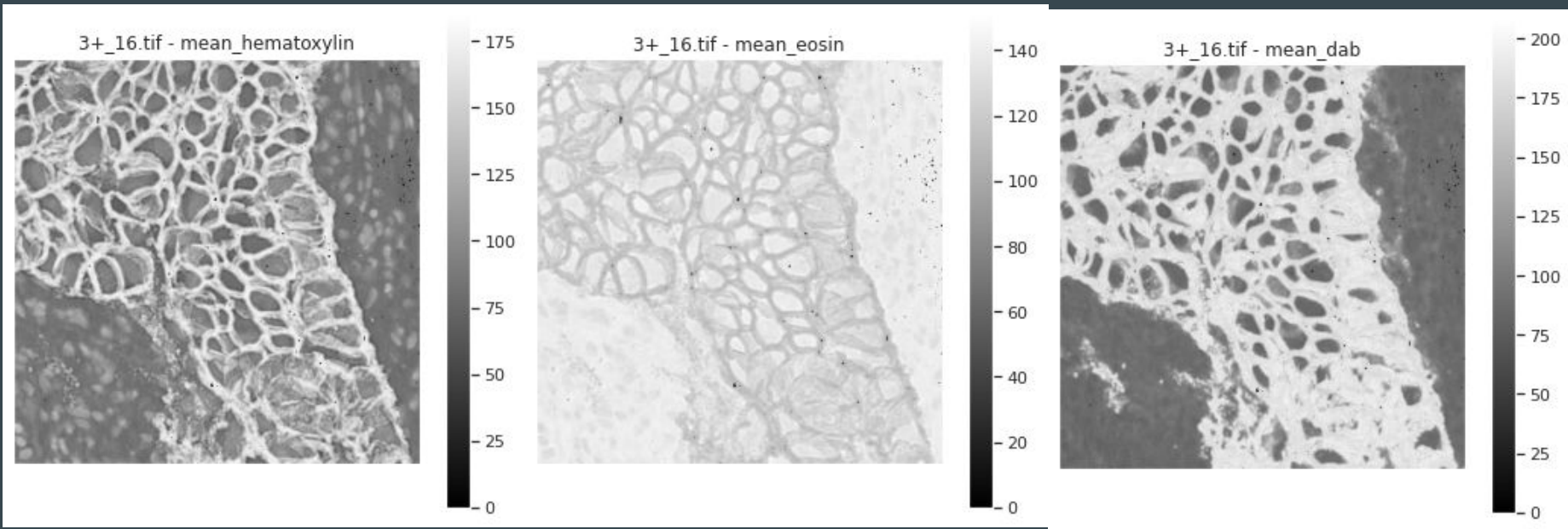
Características de Color: Gray



Extracción de Características

Extracción de
características

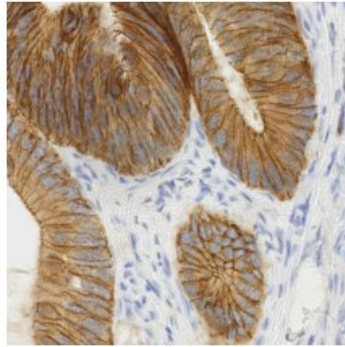
Características de Color: Haematoxylin-Eosin-DAB (HED)



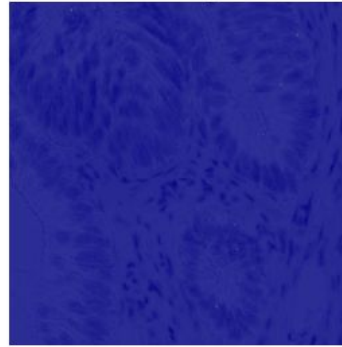
Extracción de Características

Extracción de
características

Original image



Hematoxylin



Eosin



DAB

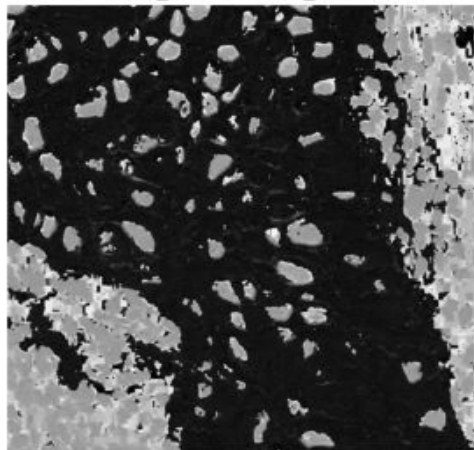


Extracción de Características

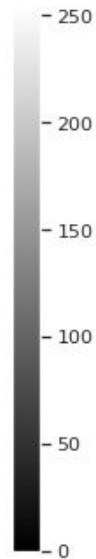
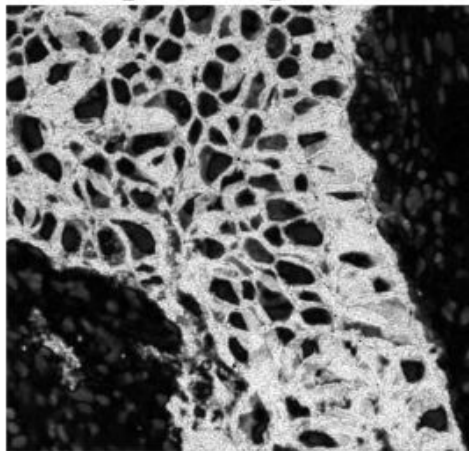
Extracción de
características

Características de Color: HSV

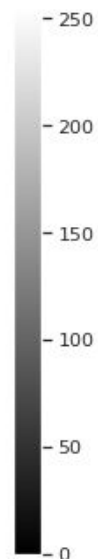
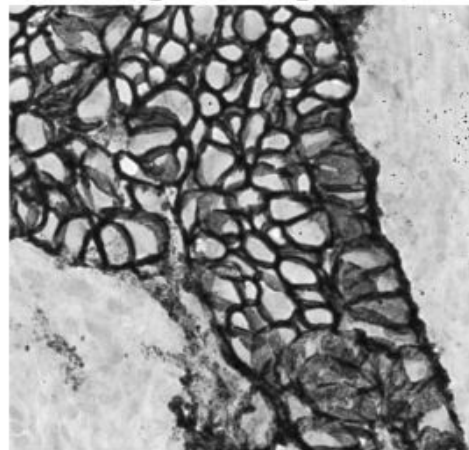
3+_16.tif - mean_hue



3+_16.tif - mean_saturation



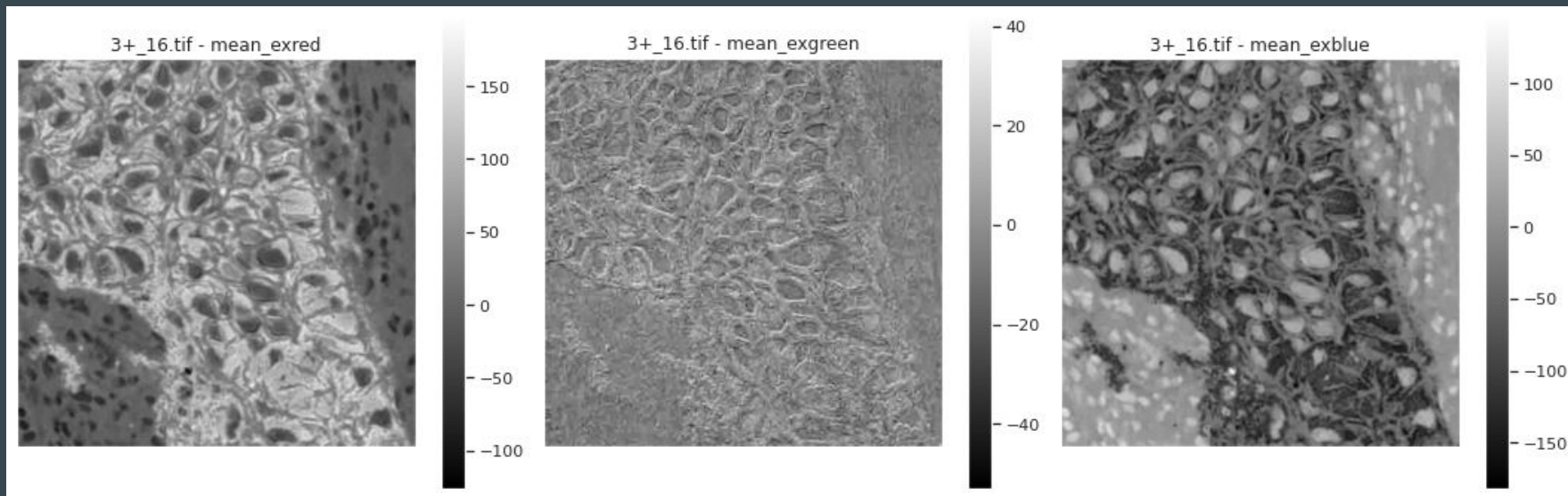
3+_16.tif - mean_value



Extracción de Características

Extracción de
características

Características de Color: RGB excess



Extracción de Características

Extracción de
características

Características de Filtro o Derivadas

Se obtienen a partir aplicar una convolución con filtros de tamaño 3x3:

- Magnitud Gradiente Sobel
- Laplacian

Extracción de Características

Extracción de
características

Características de Filtro o Derivadas: Magnitud Gradiente Sobel

$\frac{1}{8} \begin{bmatrix} -1 & 0 & 1 \\ -2 & 0 & 2 \\ -1 & 0 & 1 \end{bmatrix} \quad \frac{1}{8} \begin{bmatrix} 1 & 2 & 1 \\ 0 & 0 & 0 \\ -1 & -2 & -1 \end{bmatrix}$

$S_x \quad S_y$

$\begin{bmatrix} 1 & 0 & -1 \\ 2 & 0 & -2 \\ 1 & 0 & -1 \end{bmatrix} = \begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix} \times \begin{bmatrix} 1 & 0 & -1 \end{bmatrix}$

Sobel x-derivative

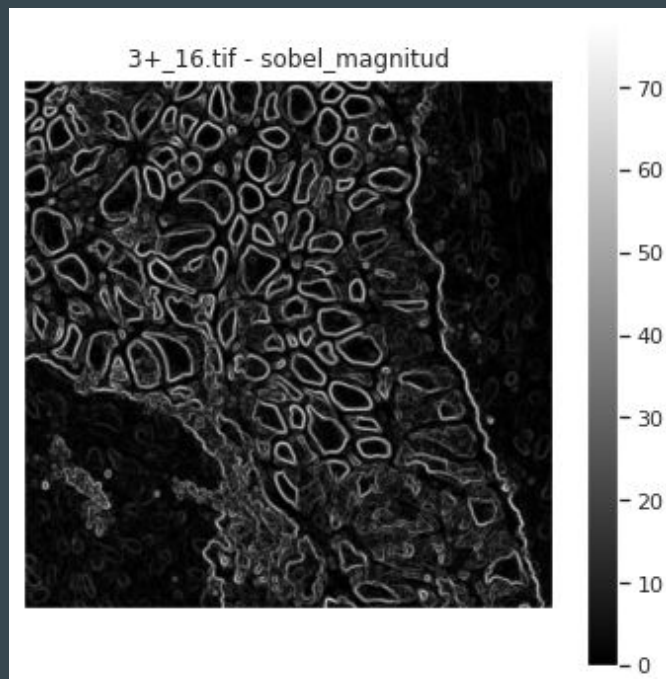
weighted average and scaling

Run filter over image

$$\frac{\partial f}{\partial x} = S_x \otimes f \quad \frac{\partial f}{\partial y} = S_y \otimes f$$

Image gradient

$$\nabla f = \left[\frac{\partial f}{\partial x}, \frac{\partial f}{\partial y} \right]$$

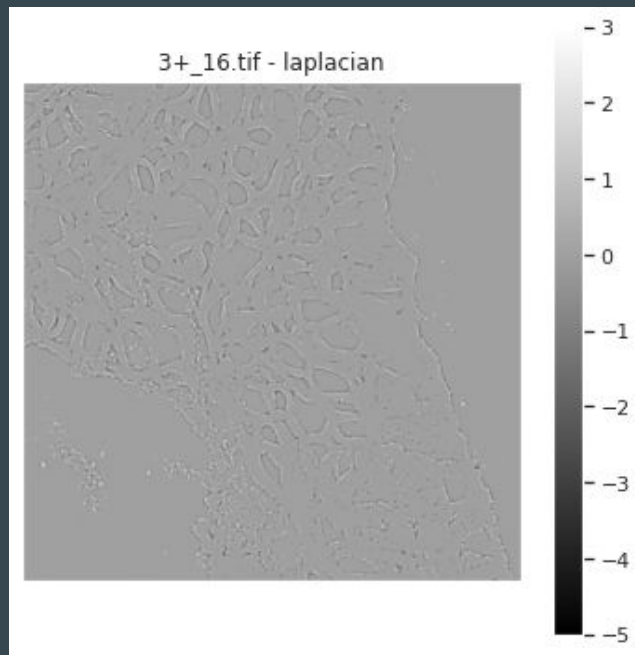


Extracción de Características

Extracción de
características

Características de Filtro o Derivadas: Laplacian

-1	-1	-1
-1	8	-1
-1	-1	-1



Extracción de Características

Extracción de
características

Características de Textura

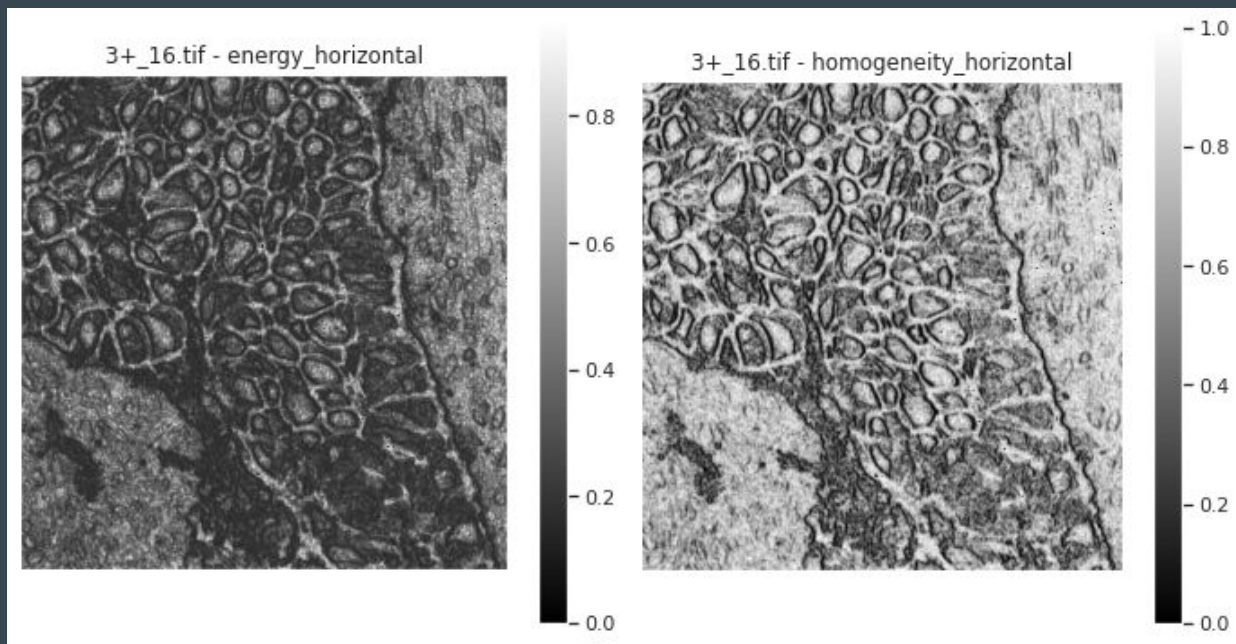
Se obtienen a partir de las características emergentes de un pequeño vecindario de píxeles de tamaño 5x5:

- GLCM (Gray Level Co-occurrences-Matrix)
 - Energy
 - Homogeneity
 - Correlation
 - Mean
- Local Binary Pattern with rotation invariant extension

Extracción de Características

Extracción de
características

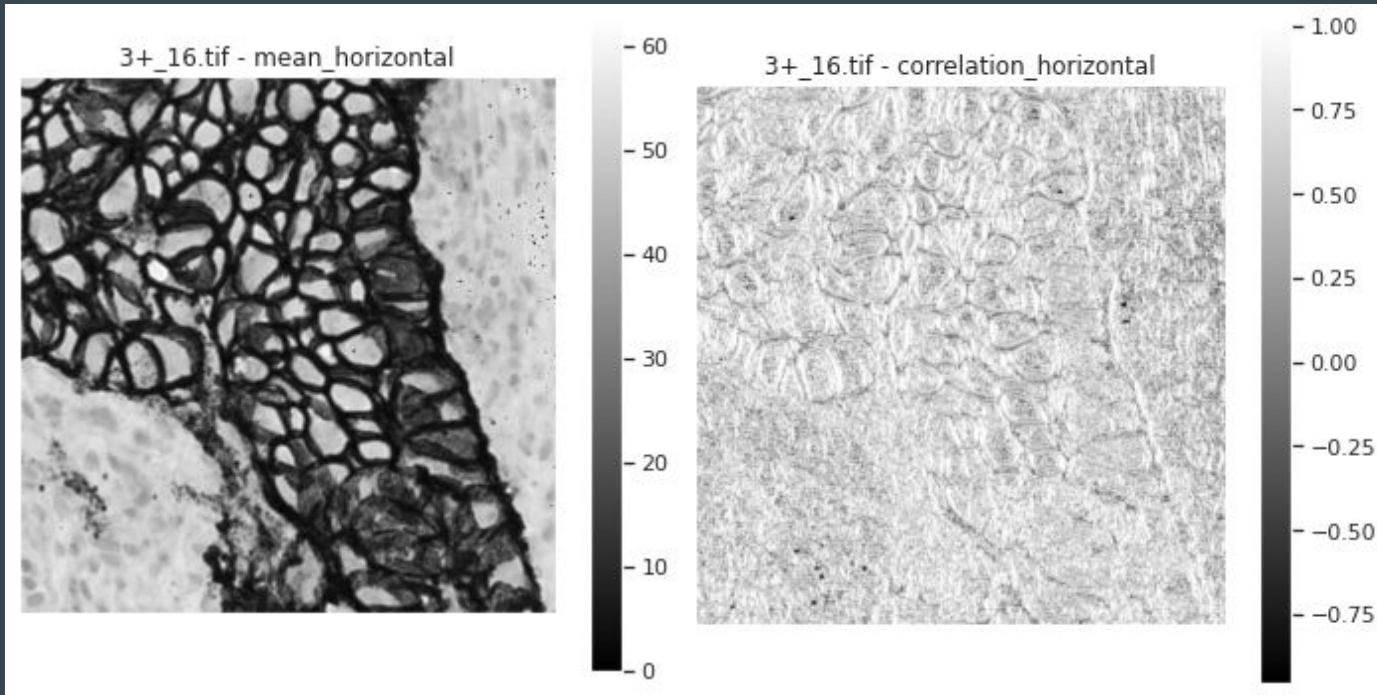
Características de Textura: GLCM



Extracción de Características

Extracción de
características

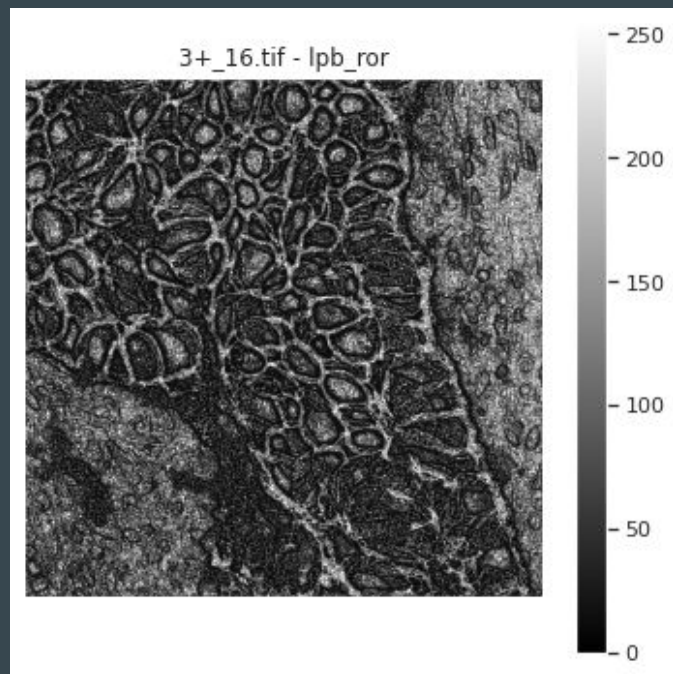
Características de Textura: GLCM



Extracción de Características

Extracción de
características

Características de Textura: LBP with rotation invariant



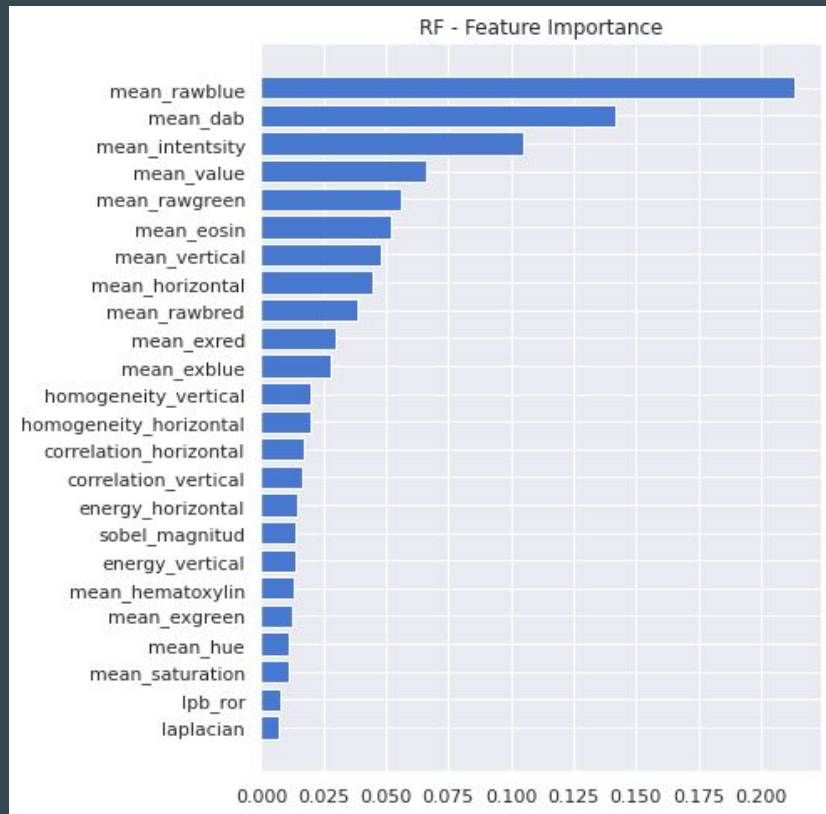
Selección de Características

Importancia de Características

Selected Features

mean_rawblue
mean_dab
mean_intentsity,
mean_rawgreen,
mean_eosin
mean_vertical
mean_rawred
homogeneity_vertical
mean_hematoxylin
sobel_magnitud

Extracción de características



Fuzzificación

Fuzzificación

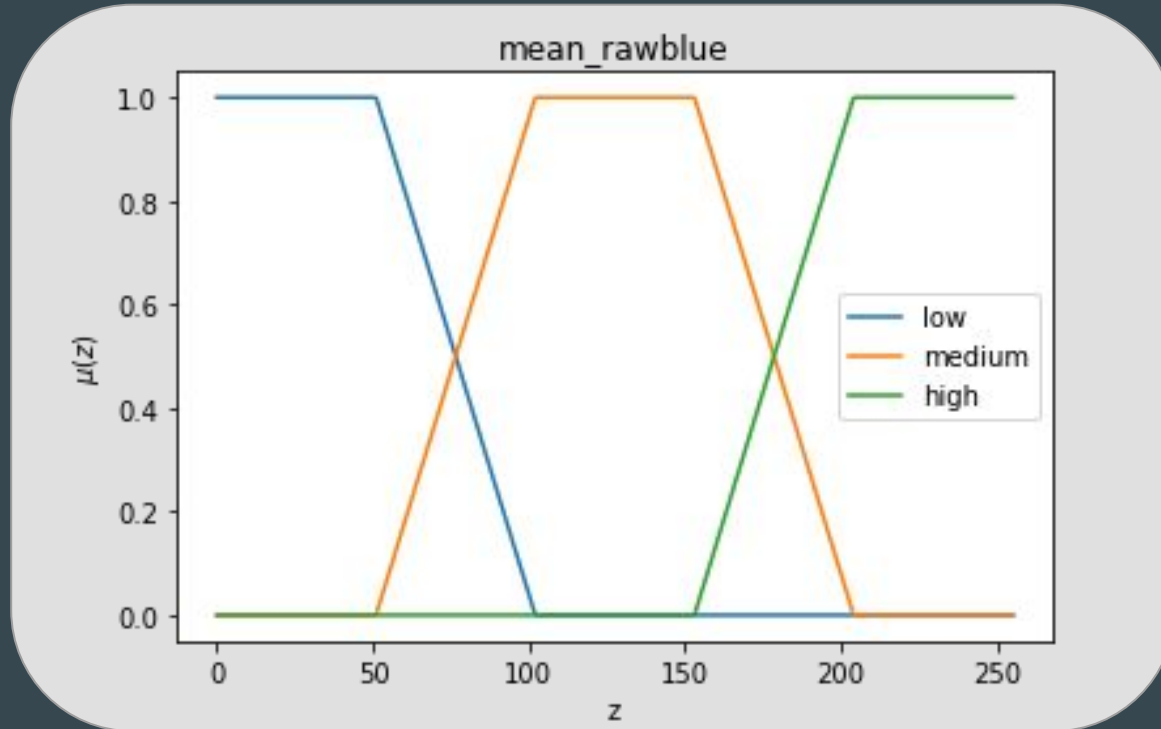
El conocimiento del dominio nos permite definir las funciones de pertenencias, por ejemplo:

- El grado de intensidad de la tinción café: {bajo, moderado, alto}

Para estos casos, se utilizarán funciones de pertenencia triangulares, trapezoidales y sigmas.

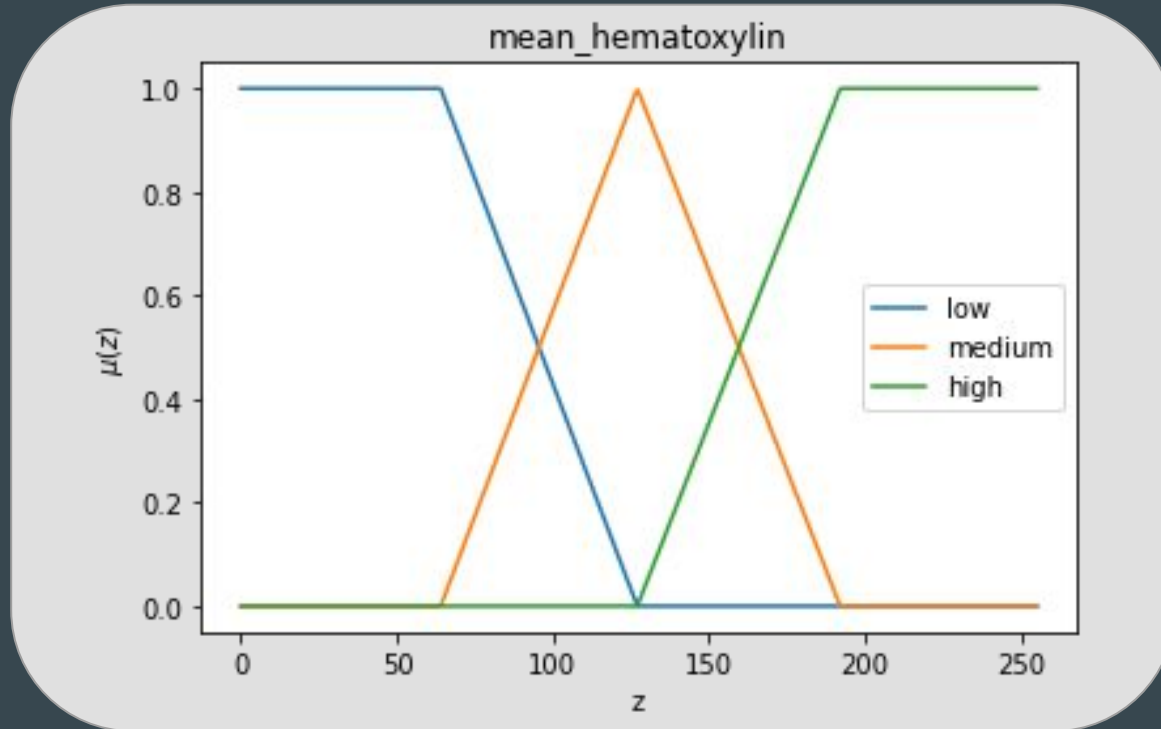
Fuzzificación

Fuzzificación



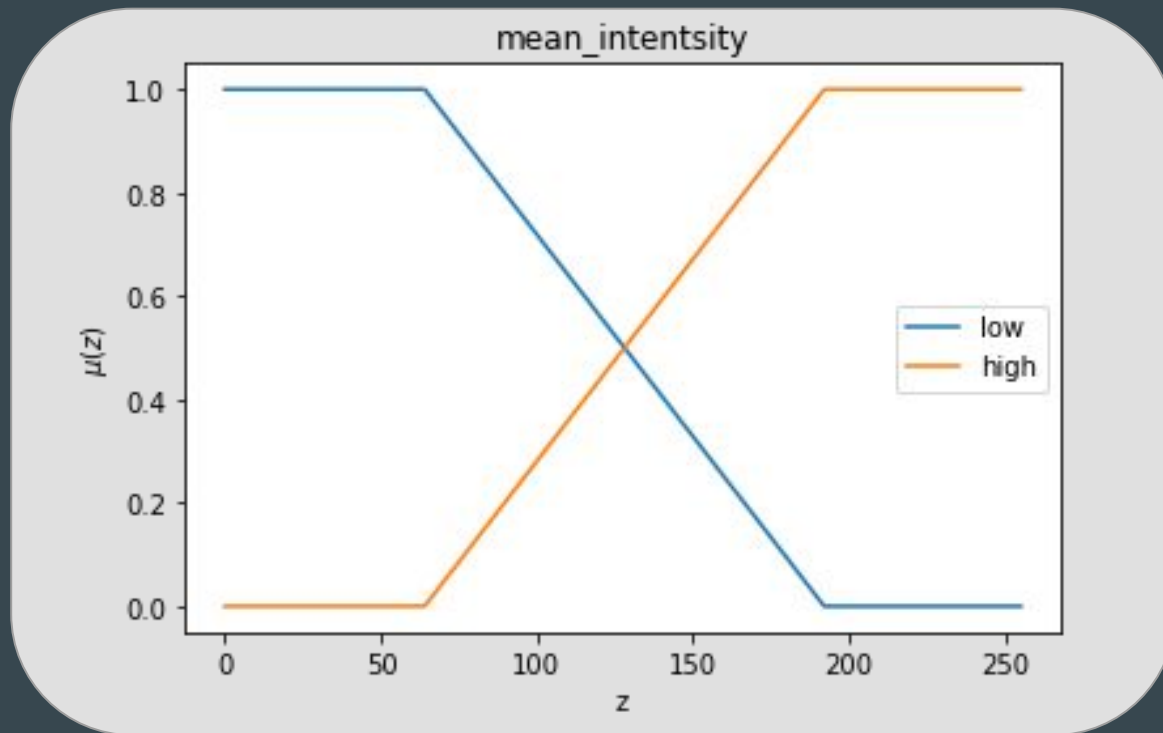
Fuzzificación

Fuzzificación



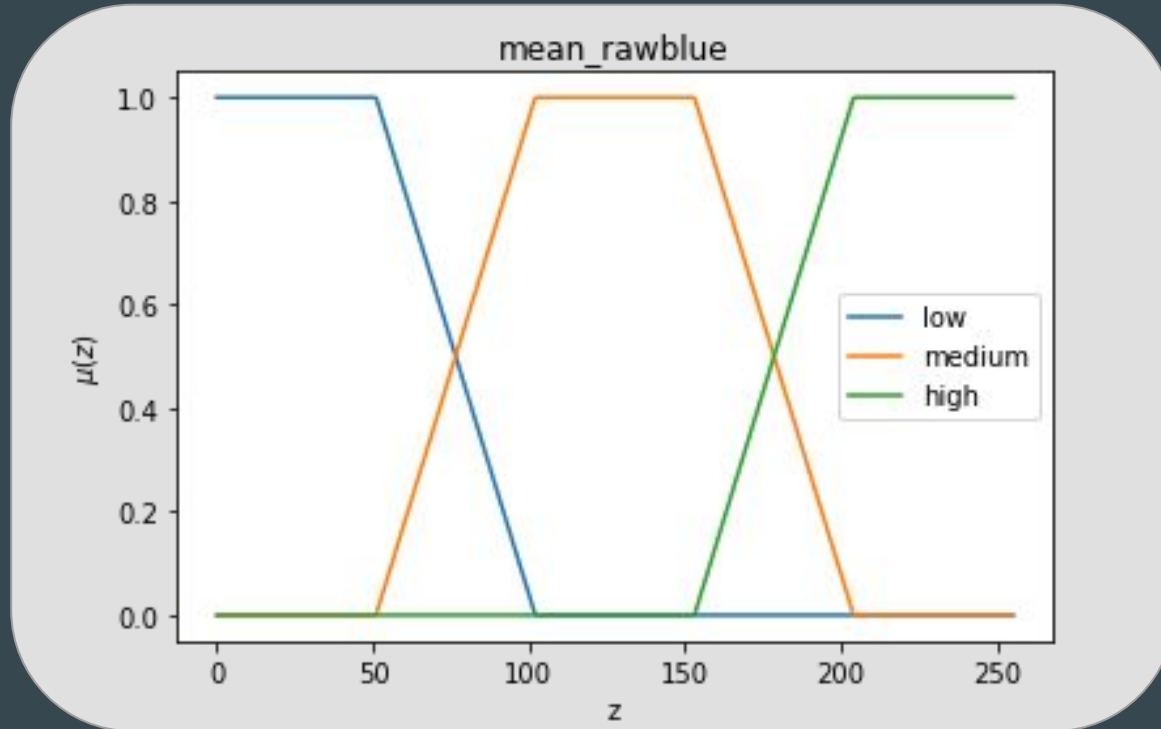
Fuzzificación

Fuzzificación



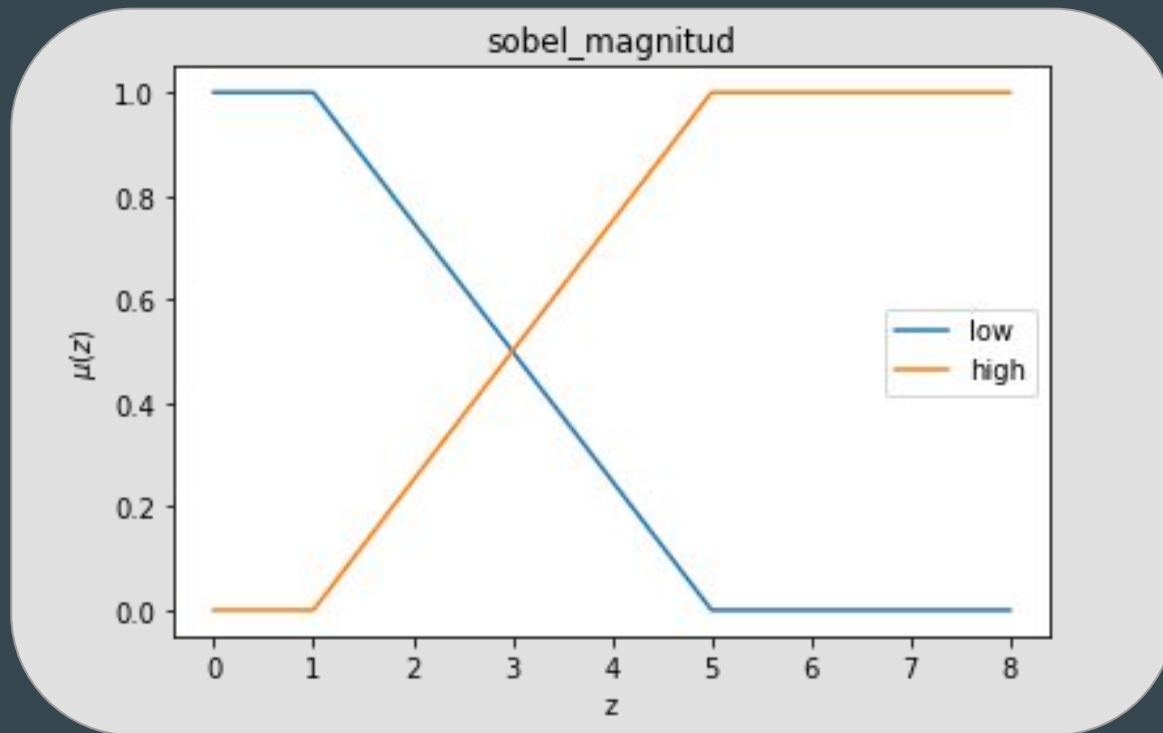
Fuzzificación

Fuzzificación



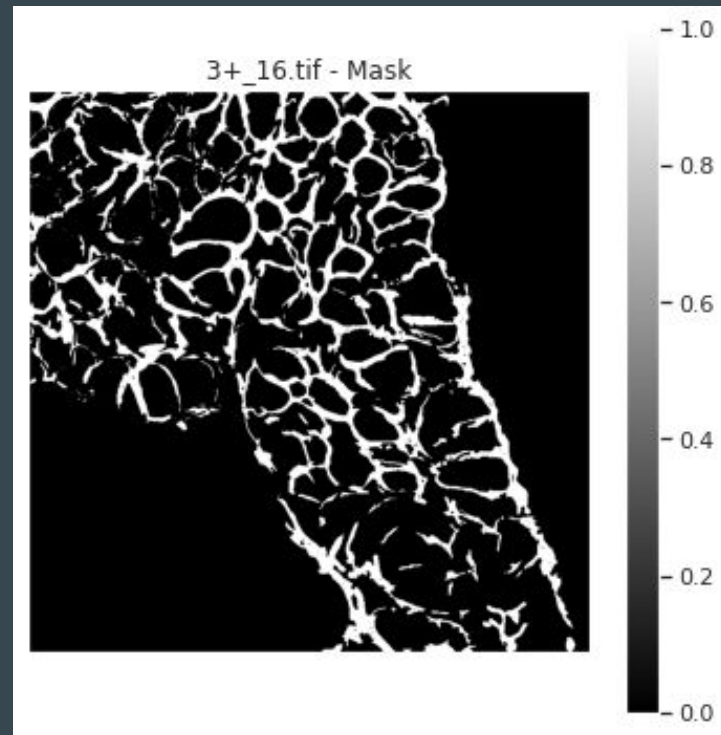
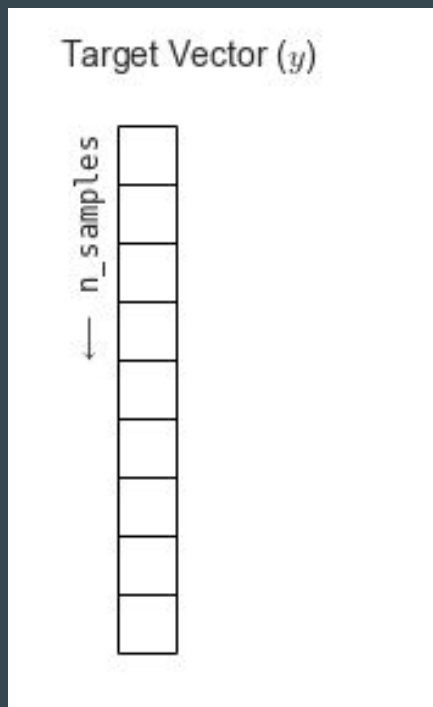
Fuzzificación

Fuzzificación



Post Processing

PostProcessing



Entrenamiento

Fuzzy Decision Tree

Fuzzy Tree 1

Considera desbalance
max_deep = 4,
theta_r = 0.98,
theta_n = 15

Fuzzy Tree 2

No considera
desbalance
max_deep = 4,
theta_r = 0.98,
theta_n = 15

Evaluación

Métrica de desempeño

Se utilizará la medida de ***Precision, Recall y F1-Score*** para determinar el desempeño de la clasificación de cada pixel.

TP: True Positive

TN: True Negative

FP: False Positivo

FN: False negative

$$precision = \frac{TP}{TP + FP}$$

$$recall = \frac{TP}{TP + FN}$$

$$F1 = \frac{2 \times precision \times recall}{precision + recall}$$

$$accuracy = \frac{TP + TN}{TP + FN + TN + FP}$$

Métrica de desempeño

Fuzzy Decision Tree

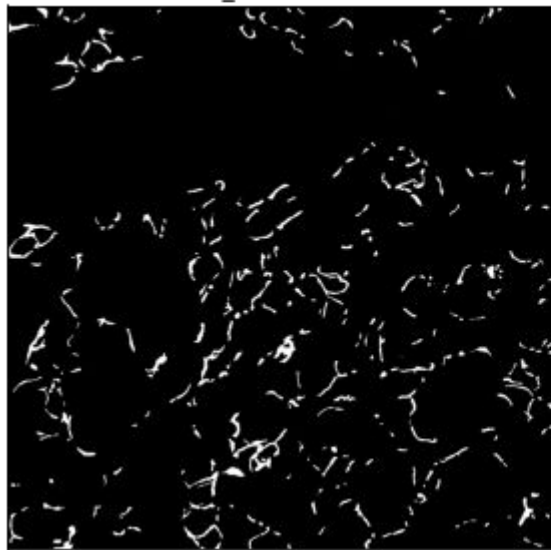
	precision	recall	f1-score	support
non-overexpression	0.99	0.88	0.93	2781471
overexpression	0.37	0.93	0.53	211306

Fuzzy Decision Tree - Imbalanced

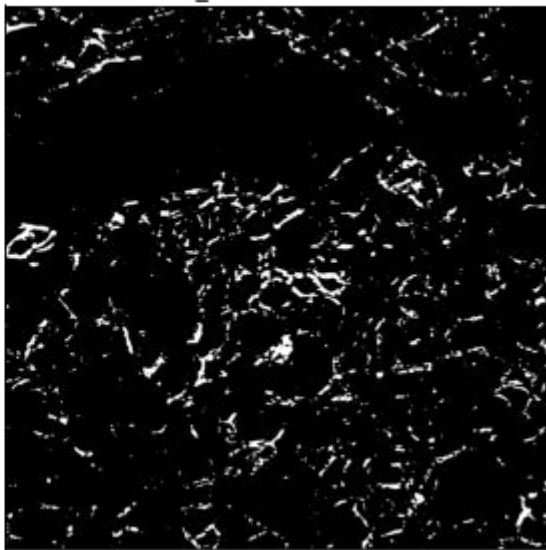
	precision	recall	f1-score	support
non-overexpression	0.95	0.99	0.97	2781471
overexpression	0.77	0.25	0.38	211306

Métrica de desempeño

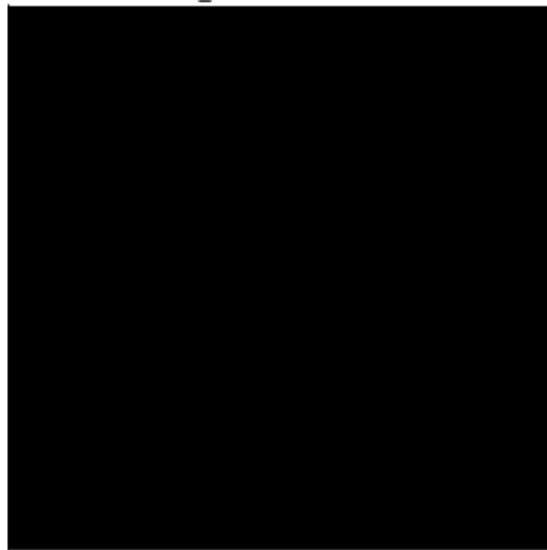
2+_9.tif - Mask



2+_9.tif - Predicted



2+_9.tif - Predicted



Métrica de desempeño

Decision Tree

	precision	recall	f1-score	support
non-overexpression	0.93	1.00	0.96	2781471
overexpression	0.97	0.04	0.08	211306

Random Forest

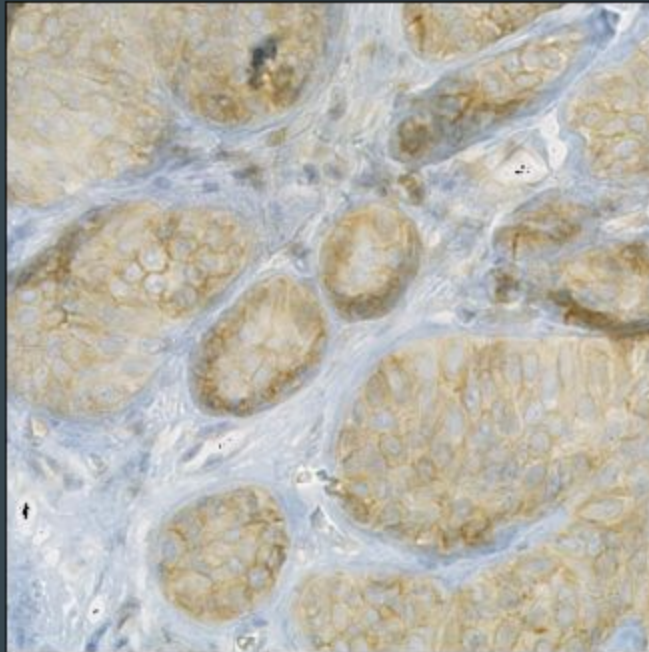
	precision	recall	f1-score	support
non-overexpression	0.99	0.92	0.95	2781471
overexpression	0.46	0.87	0.60	211306

Fuzzy Decision Tree

	precision	recall	f1-score	support
non-overexpression	0.99	0.88	0.93	2781471
overexpression	0.37	0.93	0.53	211306

Inspección Visual

1+_25.tif

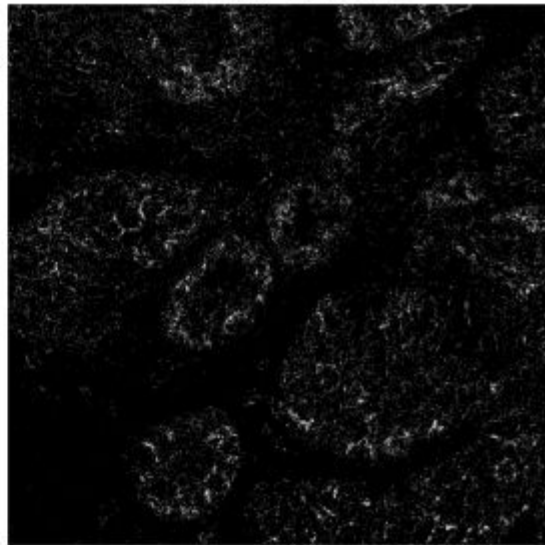


Inspección Visual: DT

1+_25.tif - Mask



1+_25.tif - Predicted

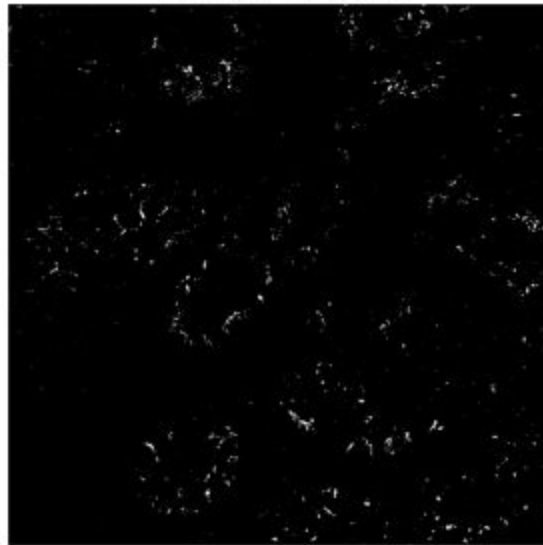


Inspección Visual: RF

1+_25.tif - Mask



1+_25.tif - Predicted

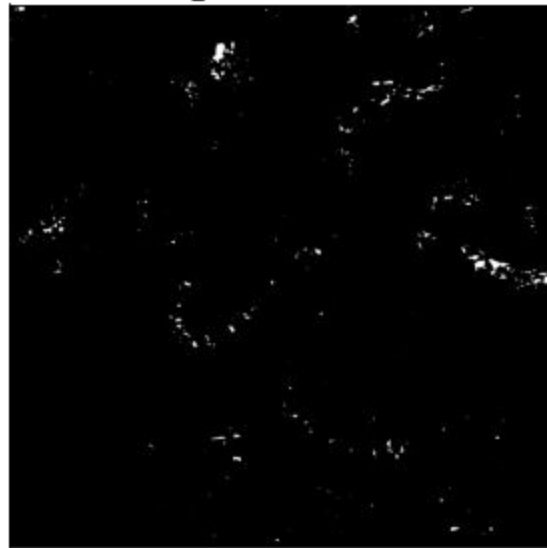


Inspección Visual: FDT

1+_25.tif - Mask

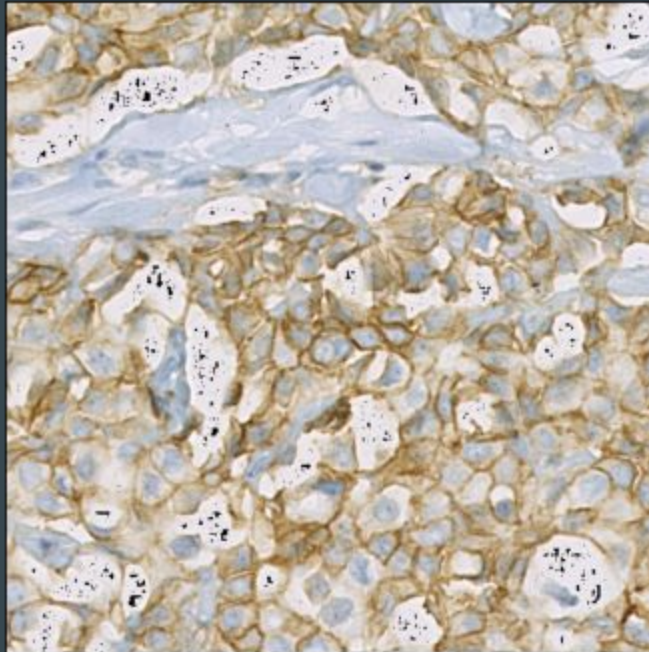


1+_25.tif - Predicted



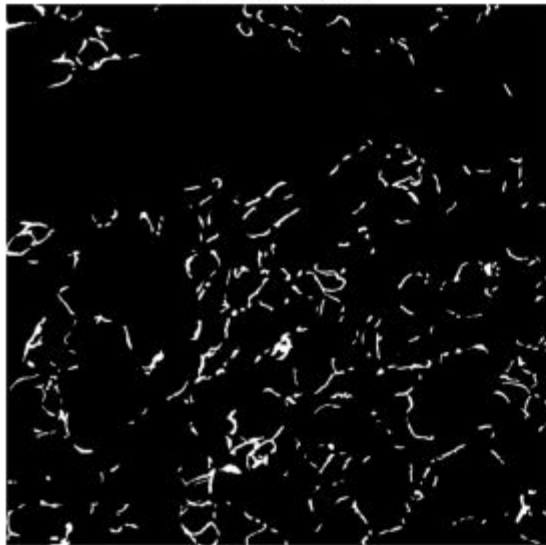
Inspección Visual

2+_9.tif

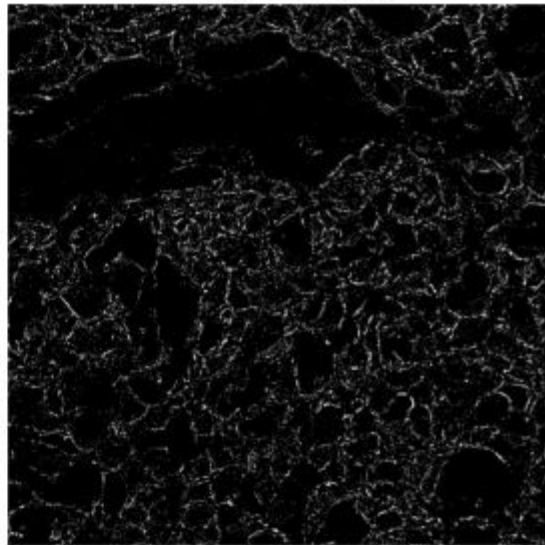


Inspección Visual: DT

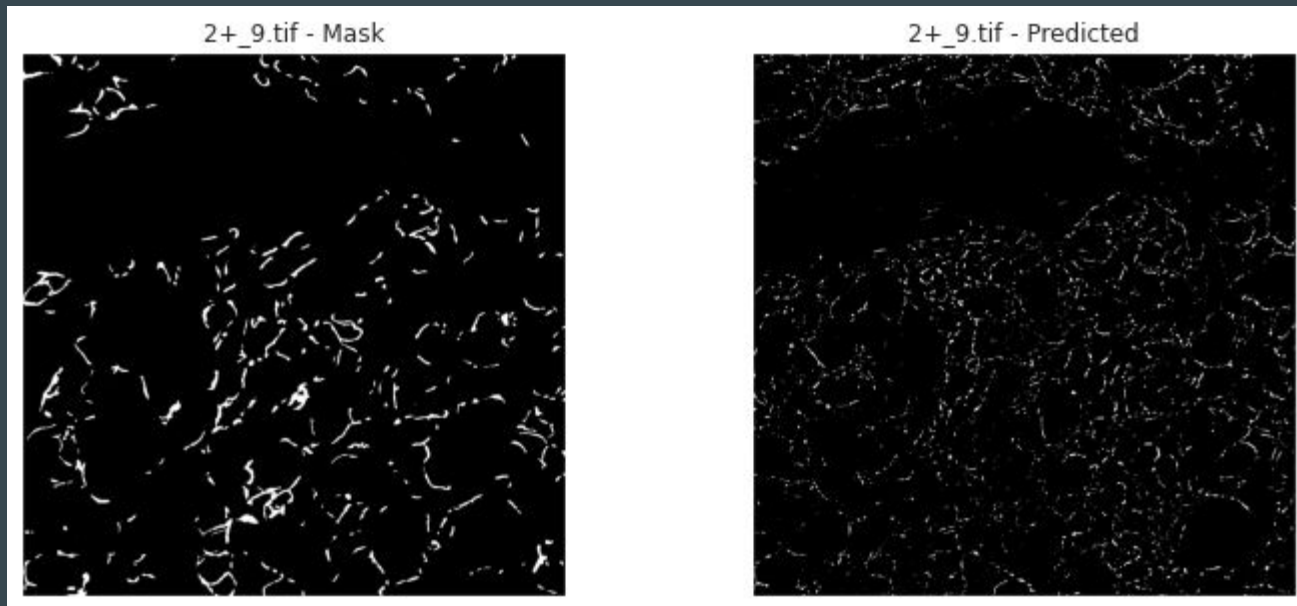
2+_9.tif - Mask



2+_9.tif - Predicted

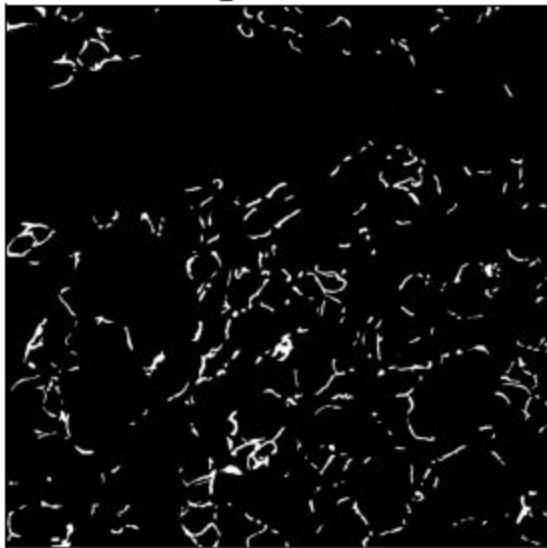


Inspección Visual: RF

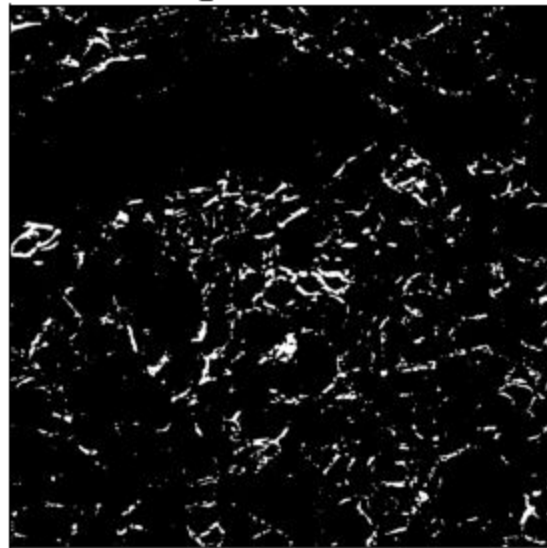


Inspección Visual: FDT

2+_9.tif - Mask

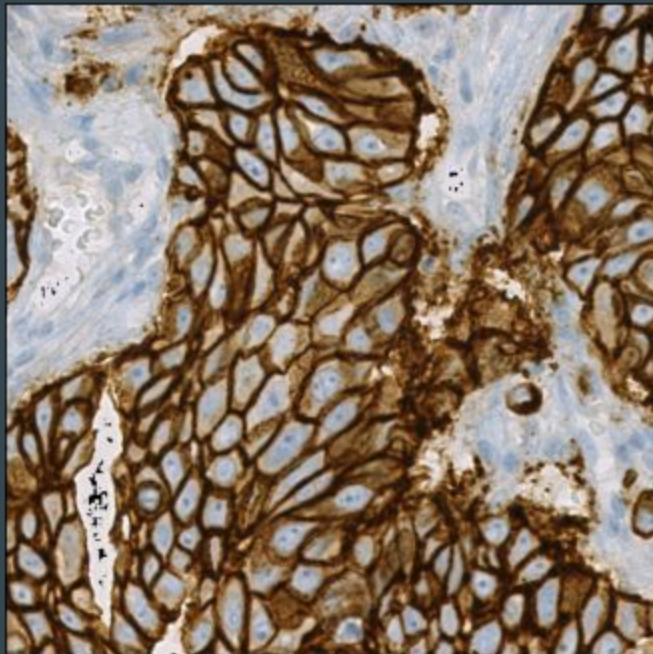


2+_9.tif - Predicted

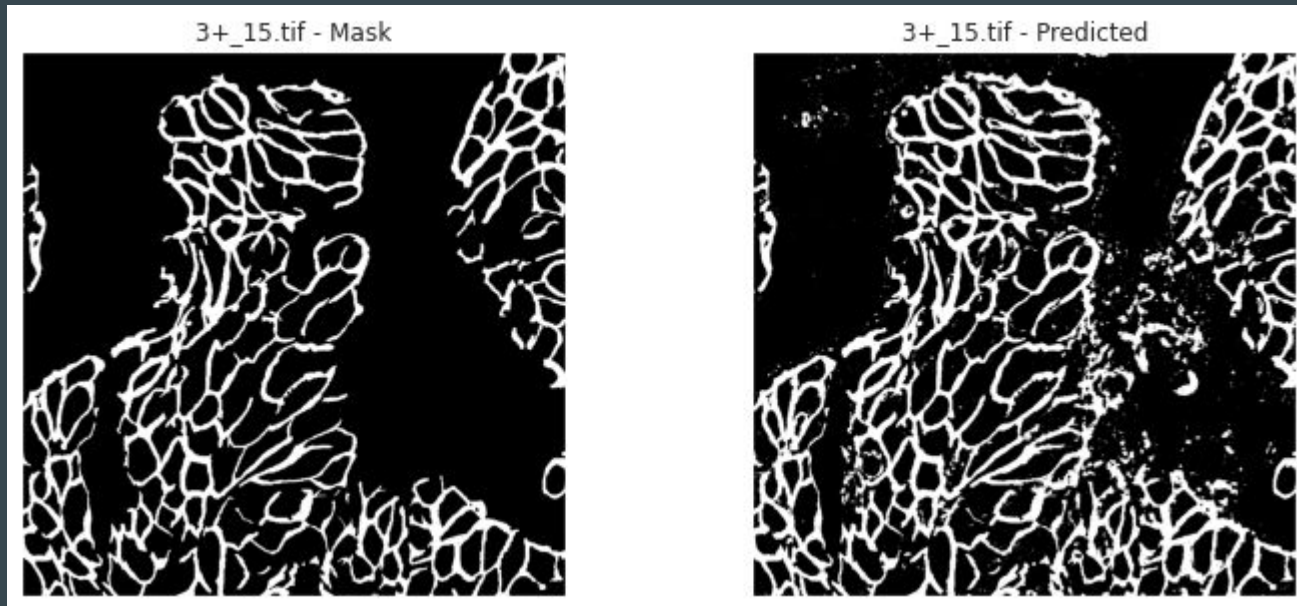


Inspección Visual

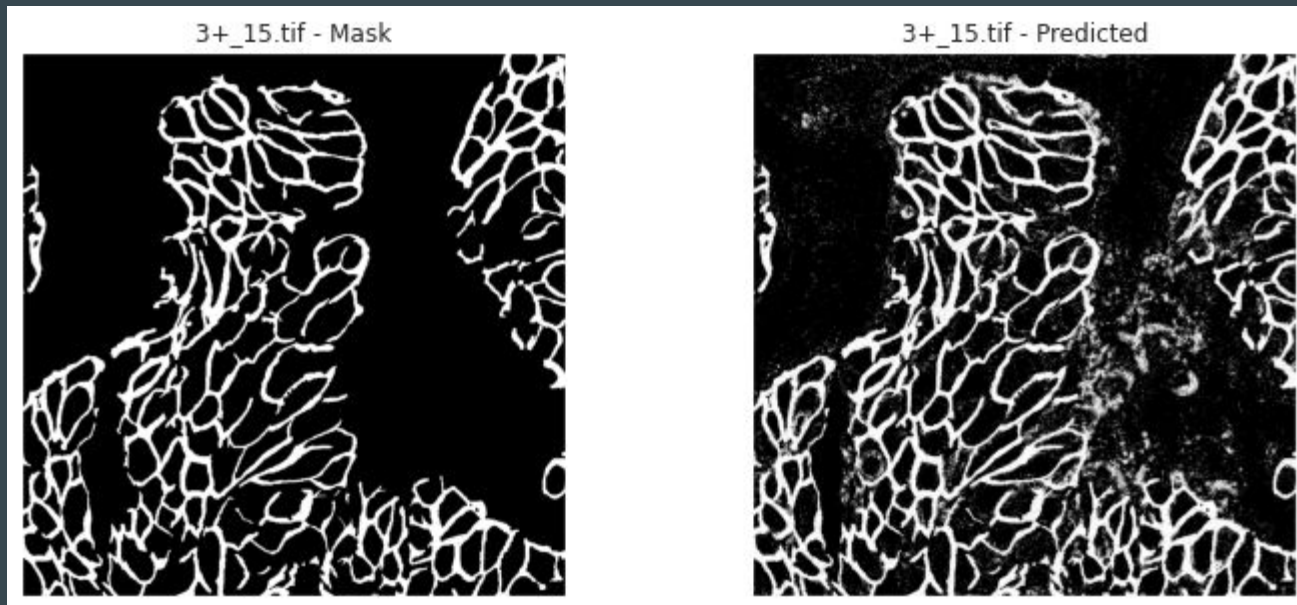
3+_15.tif



Inspección Visual: DT



Inspección Visual: RF

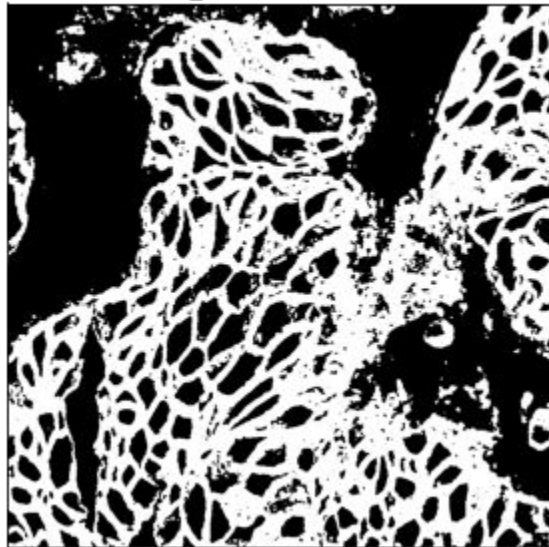


Inspección Visual: FDT

3+_15.tif - Mask



3+_15.tif - Predicted



Rule Sets: 46 Reglas para OverExpressed

Utilizando cada nodo como una condición, se agregan nuevas condiciones a medida que se desciende.

Al llegar a la hoja, se decide la clase.

Ejemplos:

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

IF ("mean_rawgreen" IS "medium") AND ("mean_rawblue" IS "low") AND ("mean_hematoxylin" IS "high") AND ("homogeneity_vertical" IS "medium") AND ("sobel_magnitud" IS "high")), THEN IS "overexpressed"

Conclusiones

Análisis de Resultados

- Se obtuvieron resultados satisfactorios con una mejora en el Recall respecto a los árboles de decisión clásicos y random forests.
- Considerar el desbalance de clases durante el entrenar con un árbol, mejora el desempeño del modelo.

Ventajas del uso de *Fuzzy Decision Trees*

- Visita todas las ramas del árbol
- Permite generar reglas de inferencia.
- Permite incluir conocimiento del dominio
- Alta interpretabilidad del árbol y reglas generadas.

Desventajas del uso de *Fuzzy Decision Trees*

- Visita todas las ramas del árbol, lo que implica más tiempo de cómputo.

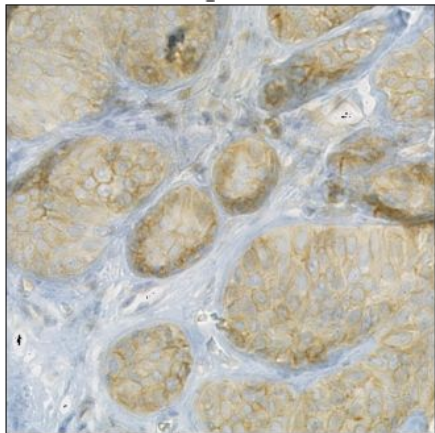
Referencias

Referencias

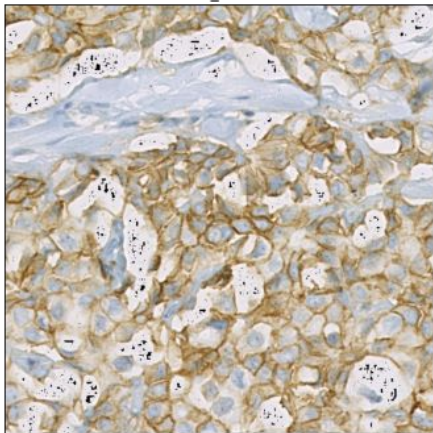
- Begelman, G., Gur, E., Rivlin, E., Rudzsky, M., & Zalevsky, Z. (2004, October). Cell nuclei segmentation using fuzzy logic engine. In 2004 International Conference on Image Processing, 2004. ICIP'04. (Vol. 5, pp. 2937-2940). IEEE.
- Hall-Beyer, M. (2000). GLCM texture: a tutorial. National Council on Geographic Information and Analysis Remote Sensing Core Curriculum, 3.
- Pezoa, R., Salinas, L., Torres, C., Härtel, S., Maureira-Fredes, C., & Arce, P. (2016, October). Segmentation of HER2 protein overexpression in immunohistochemically stained breast cancer images using Support Vector Machines. In Journal of Physics: Conference Series (Vol. 762, No. 1, pp. 528-536).
- Gonzalez, R. C., & Richard, E. (2002). Woods, digital image processing. ed: Prentice Hall Press, ISBN 0-201-18075, 8.
- Ruifrok, A. C., & Johnston, D. A. (2001). Quantification of histochemical staining by color deconvolution. Analytical and quantitative cytology and histology, 23(4), 291-299.
- Tabakov, M., & Kozak, P. (2014). Segmentation of histopathology HER2/neu images with fuzzy decision tree and Takagi–Sugeno reasoning. Computers in biology and medicine, 49, 19-29.
- Tanaka, K. (1997). An introduction to fuzzy logic for practical applications.
- Umanol, M., Okamoto, H., Hatono, I., Tamura, H. I. R. O. Y. U. K. I., Kawachi, F., Umedzu, S., & Kinoshita, J. (1994, June). Fuzzy decision trees by fuzzy ID3 algorithm and its application to diagnosis systems. In Proceedings of 1994 IEEE 3rd International Fuzzy Systems Conference (pp. 2113-2118). IEEE.

Entrada

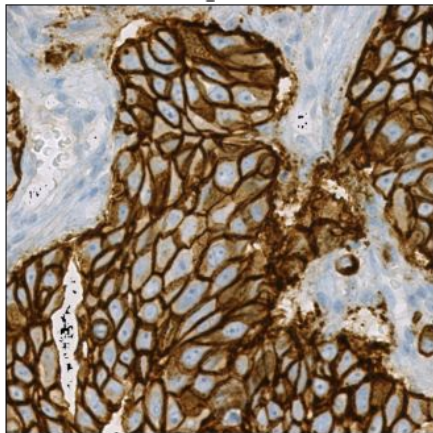
1+_25.tif



2+_9.tif



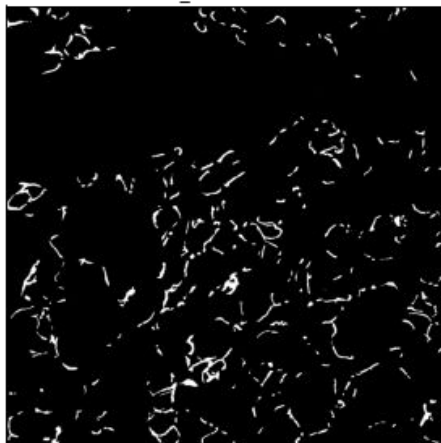
3+_15.tif



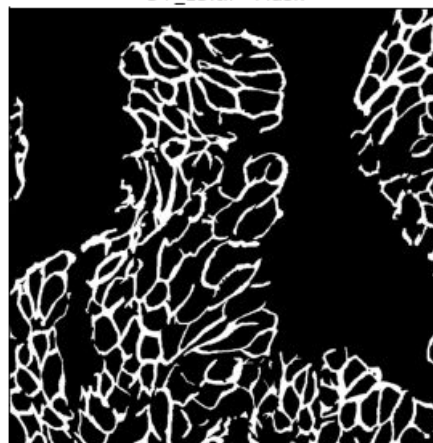
1+_25.tif - Mask



2+_9.tif - Mask



3+_15.tif - Mask



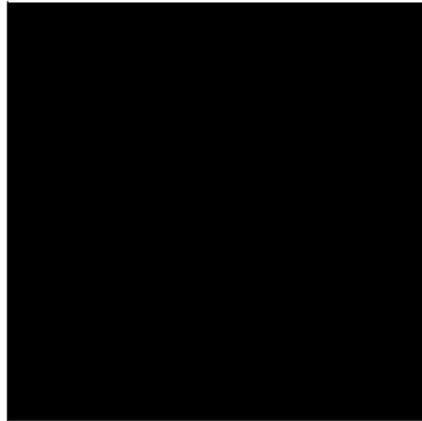
Ground Truth

FDT A

1+_25.tif - Predicted



2+_9.tif - Predicted

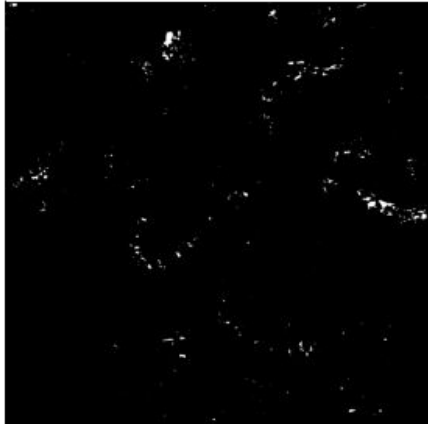


3+_15.tif - Predicted

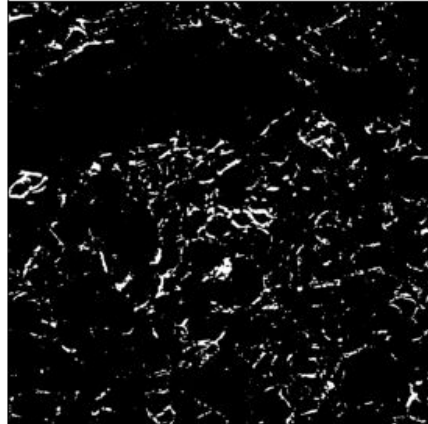


FDT B

1+_25.tif - Predicted



2+_9.tif - Predicted

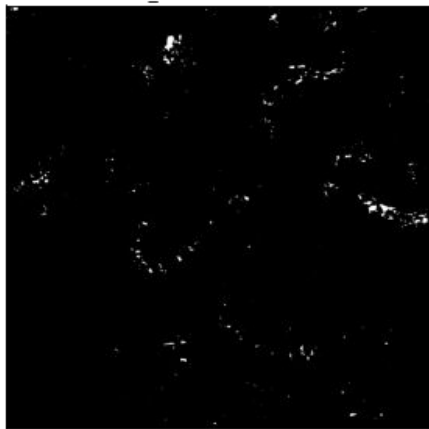


3+_15.tif - Predicted

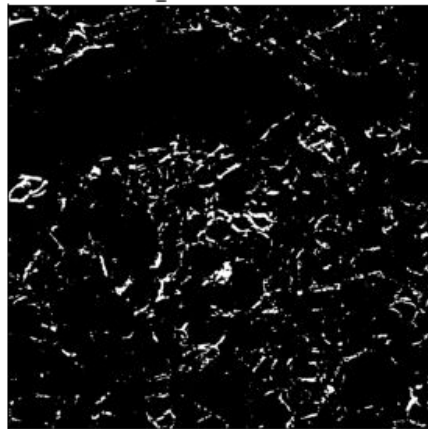


FDT C

1+_25.tif - Predicted



2+_9.tif - Predicted

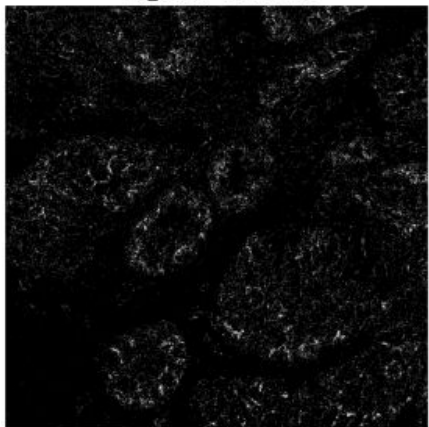


3+_15.tif - Predicted

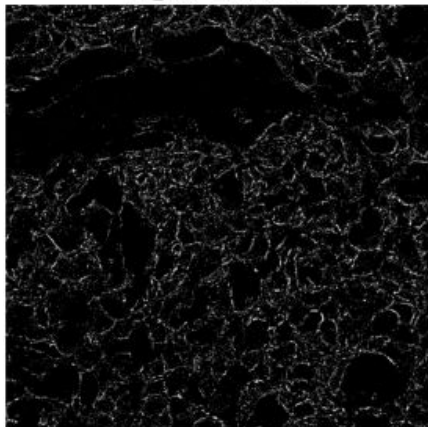


DT

1+_25.tif - Predicted



2+_9.tif - Predicted

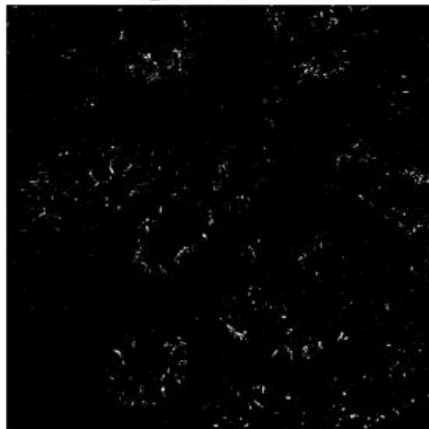


3+_15.tif - Predicted

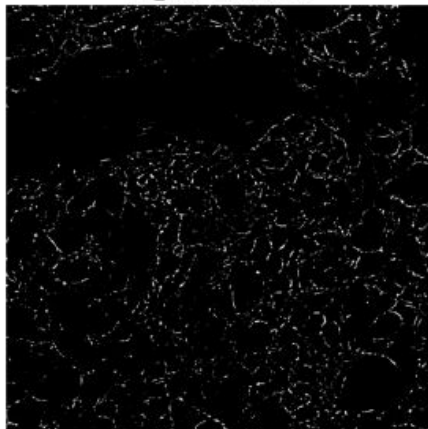


RF

1+_25.tif - Predicted



2+_9.tif - Predicted



3+_15.tif - Predicted

