

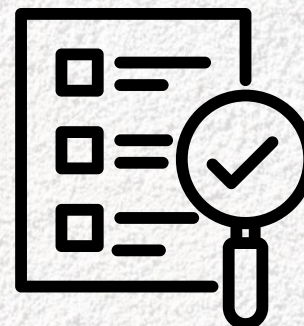
RECONOCIMIENTO DE PATRONES

ANÁLISIS Y SEGMENTACION DE MASTOCITOS EN IMAGENES DE PATOLOGÍA

García López Erik – Reyes Herrera Rogelio – Serapio Hernández Alexis Arturo

OBJETIVOS GENERALES:

- Desarrollar un sistema automatizado para la detección y clasificación de mastocitos en imágenes microscópicas, diferenciando entre células sanas y aquellas con anomalías relacionadas con mastocitomas.
- Contribuir a la medicina veterinaria mediante herramientas tecnológicas que mejoren la precisión y rapidez en el diagnóstico de mastocitomas en perros.



OBJETIVOS ESPECÍFICOS:

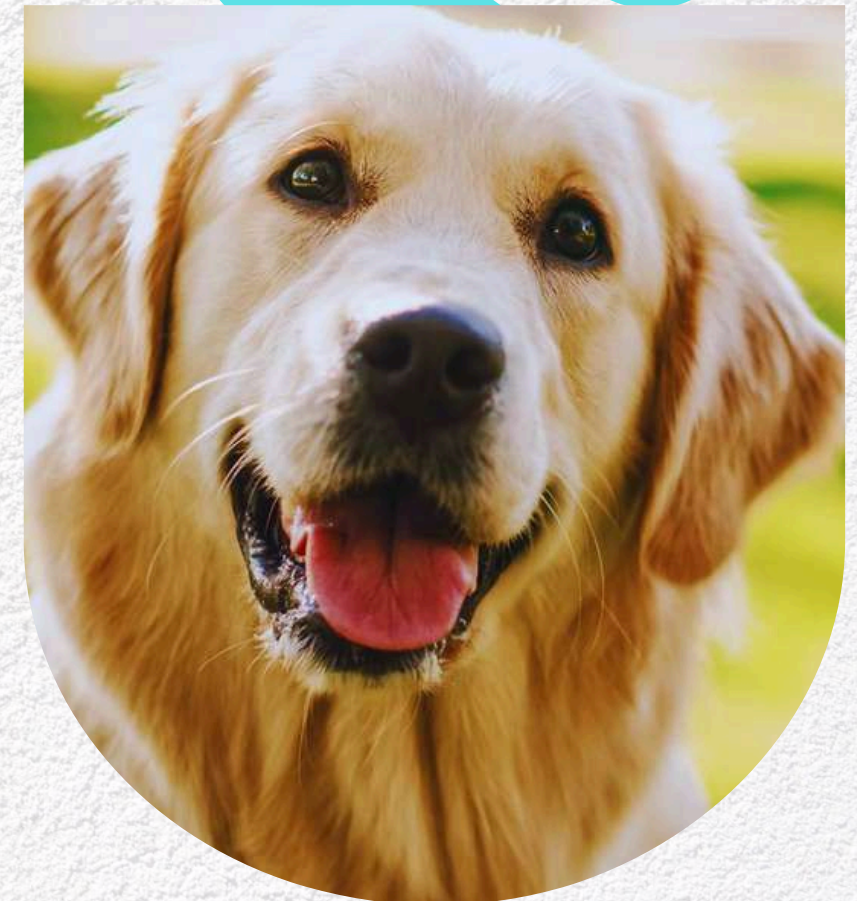
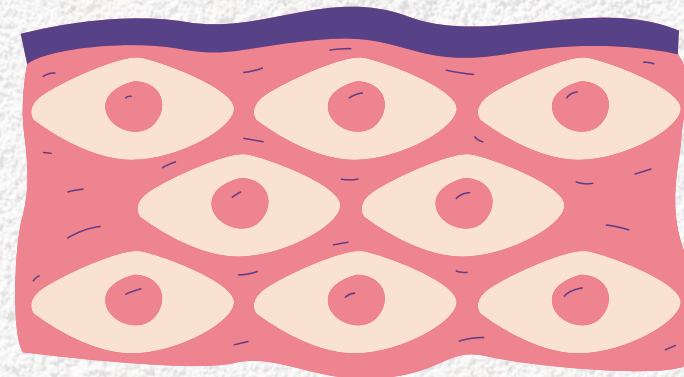
- Recolectar y procesar imágenes microscópicas de tejidos caninos
- Implementar técnicas de visión computacional, como transformadas de Hough o segmentación para aislar mastocitos.
- Identificar y caracterizar características relevantes de los mastocitos.
- Entrenar modelos de clasificación



INTRODUCCIÓN

El diagnóstico y tratamiento oportuno de enfermedades en animales domésticos representa un reto importante en la medicina veterinaria moderna.

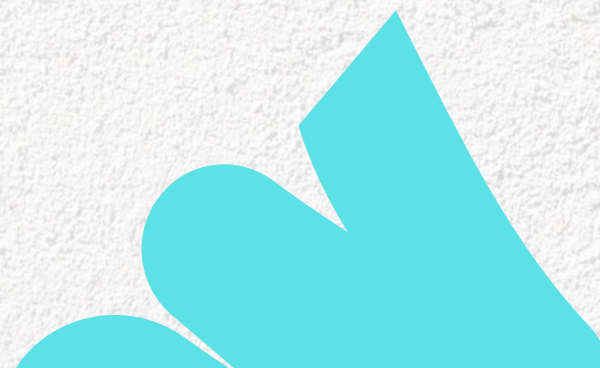
Entre estas enfermedades, los mastocitomas son un tipo común de tumor cutáneo en perros, destacando debido a su potencial agresividad y alta incidencia



INTRODUCCIÓN



- El mastocito es una célula del sistema inmunológico, reconocible por su núcleo y el citoplasma granular que lo rodea.
- Por otra parte, las células que no son mastocitos son más pequeñas y carecen de las características granulares visibles en su citoplasma.

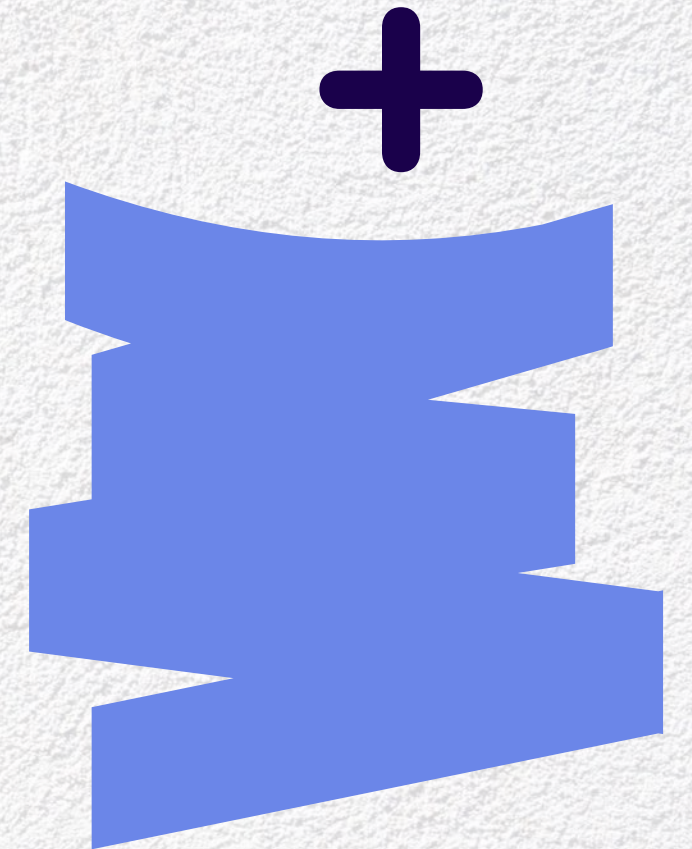


INTRODUCCIÓN

La identificación precisa de estas afecciones suele depender de un análisis histopatológico realizado por especialistas, lo que suele ser un proceso laborioso y propenso a variaciones humanas.



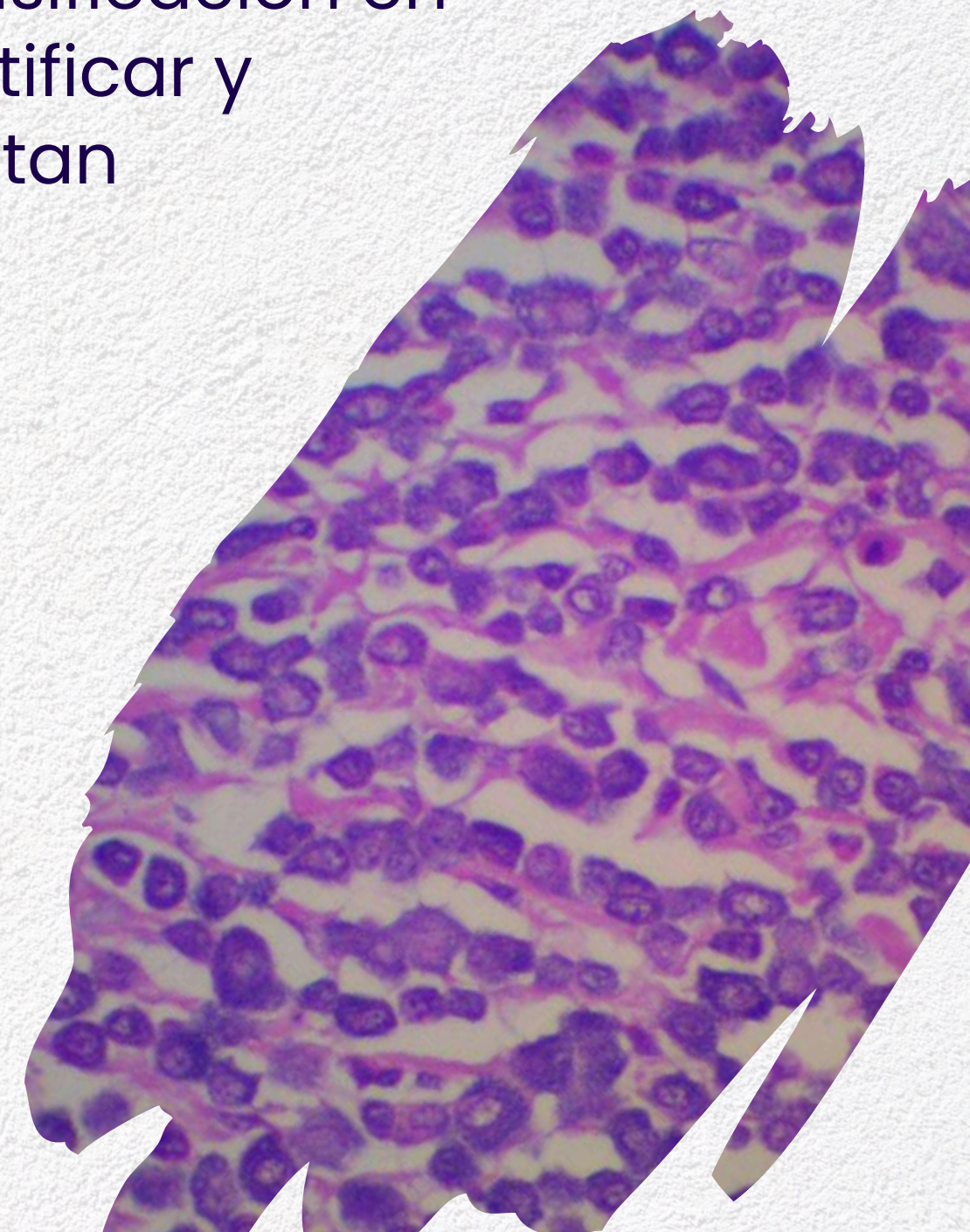
En este contexto, el uso de **técnicas de visión computacional y el reconocimiento de patrones** puede fungir como una herramienta prometedora para automatizar y mejorar la precisión del diagnóstico.



INTRODUCCIÓN

Por lo que este proyecto, busca aplicar métodos de clasificación en imágenes microscópicas de tejidos que permitan identificar y clasificar mastocitos saludables y aquellos que presentan anomalías relacionadas con mastocitomas.

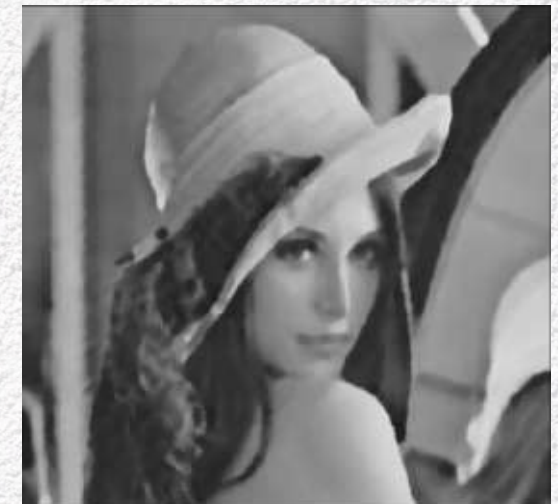
Esto nos **permite extraer información cuantitativa sobre las propiedades visuales de las células**, como patrones de distribución, contraste, homogeneidad y rugosidad para detectar las regiones de mastocitos y determinar la situación en la que se encontraría un animalito.



INTRODUCCIÓN

Se utilizaron diversas tecnicas como:

- Etiquetado manual con labelme
- Conversión a escala de grises
- Suavizado mediante filtro gaussiano
- Detección de bordes con el algoritmo Canny.
- GLCM
- SVM
- K means (no se formalizó)



METODOLOGÍA

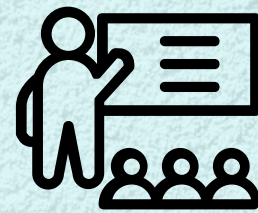
1

**Recolectar la
información**



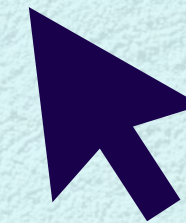
2

**Elegir el modelo
adecuado**



3

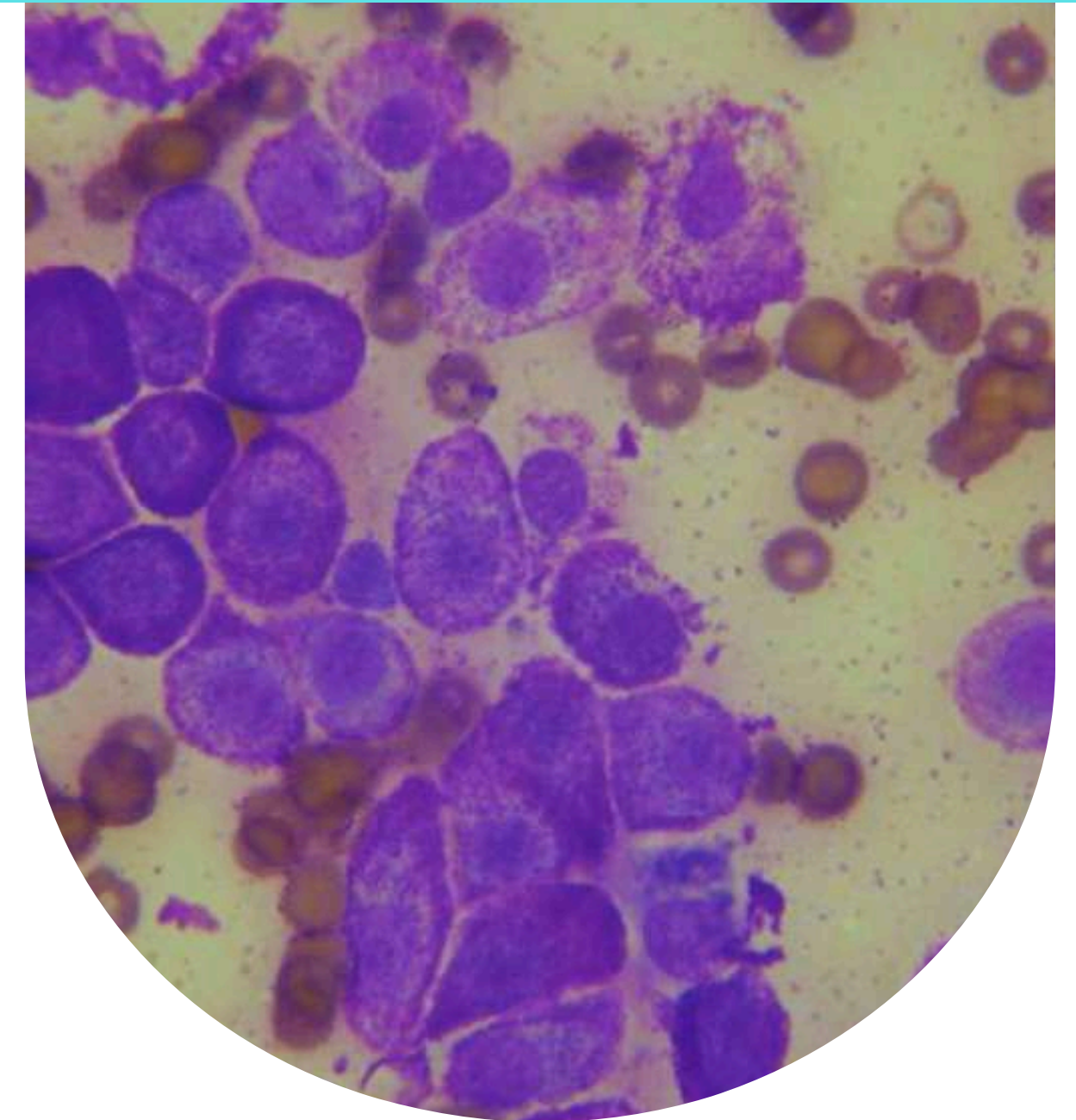
Detección



¿CÓMO SE EMPEZÓ A MANEJAR EL PROBLEMA?

Para dar solución al problema planteado en el proyecto primero había que pensar en cuáles eran las opciones más óptimas.

Al principio se desconocía la naturaleza de los datos así que pensamos en un K-means con 2 clusters.

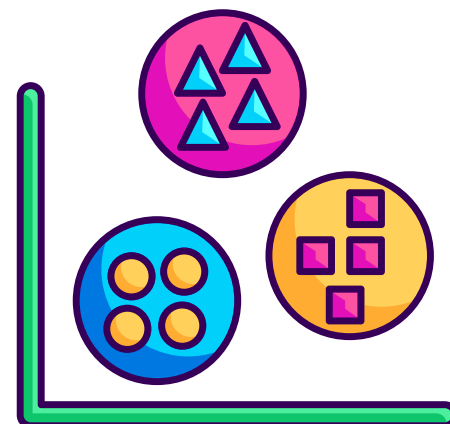


¿CÓMO SE EMPEZÓ A MANEJAR EL PROBLEMA?

El objetivo principal del algoritmo K-Means es minimizar la variabilidad dentro de cada grupo y maximizar la separación entre los grupos. Lo logra **distribuyendo los datos en k clusters**, donde **k** es un parámetro definido por el usuario.

En este caso aplicamos 2 clústers, mastocitos sanos y mastocitomas.

Pero al ser células tan parecidas hasta cierto **punto clasificaba 50-50** como si las clases estuvieran perfectamente balanceadas; por lo tanto para recabar la información restante **consultamos directamente con el Dr. Jacobo.**



RECOLECCION: CARACTERISTICAS

Jacobo nos explicó las diferencias clave entre mastocitos sanos y otros elementos.

- **MASTOCITOS**

El núcleo suele ser pequeño, redondeado y central, con cromatina bien organizada.

Contienen **abundantes gránulos citoplasmáticos metacromáticos** que se tiñen intensamente con tinciones específicas como Giemsa o Azul de Toluidina. **Estos gránulos suelen estar distribuidos uniformemente.**

Hay un índice mitótico muy bajo o ausente, ya que estas células están maduras y no suelen estar en división activa.

- **MASTOCITOMAS**

Pueden tener núcleos grandes, multinucleación o formas nucleares atípicas, como núcleos alargados o plegados.

Los gránulos pueden ser escasos, irregulares, o incluso ausentes en algunos casos. Además, la calidad de la tinción puede verse alterada, con pérdida de metacromasia en los gránulos.

Muestran un índice mitótico elevado, reflejando actividad proliferativa asociada con procesos tumorales.

RECOLECCIÓN

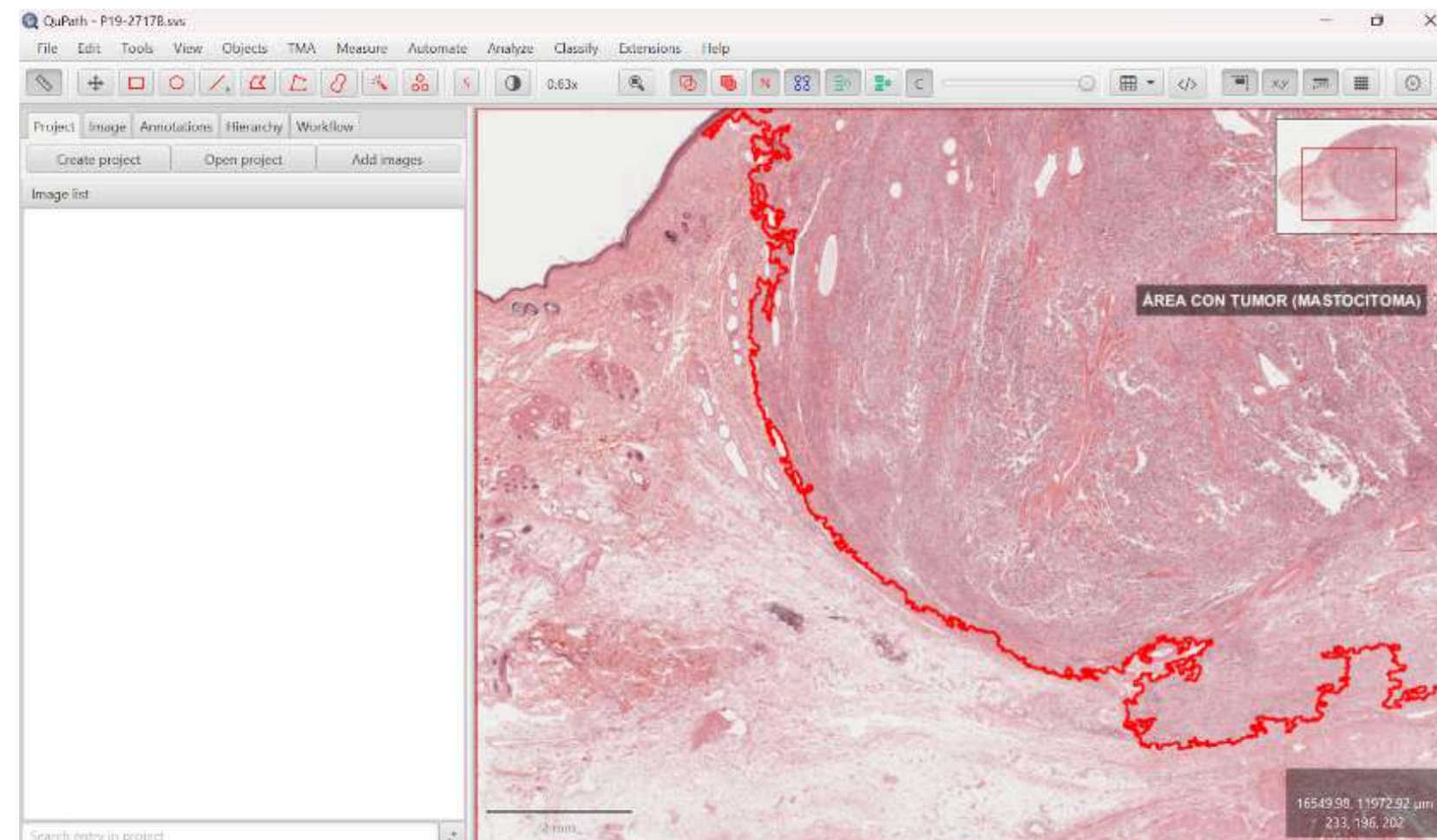
Parte importante del proceso de reconocimiento de patrones es contar con la información suficiente y con características que sirvan para encontrar información relevante en los datos que no podrían verse a simple vista o que sería muy tardado hacer.

Para ello, de la imagen proporcionada, buscamos zonas que representaran ambas clases de células y etiquetamos manualmente al rededor de 233 células. Esto gracias a la ayuda del Dr. Jacobo Carrizosa.



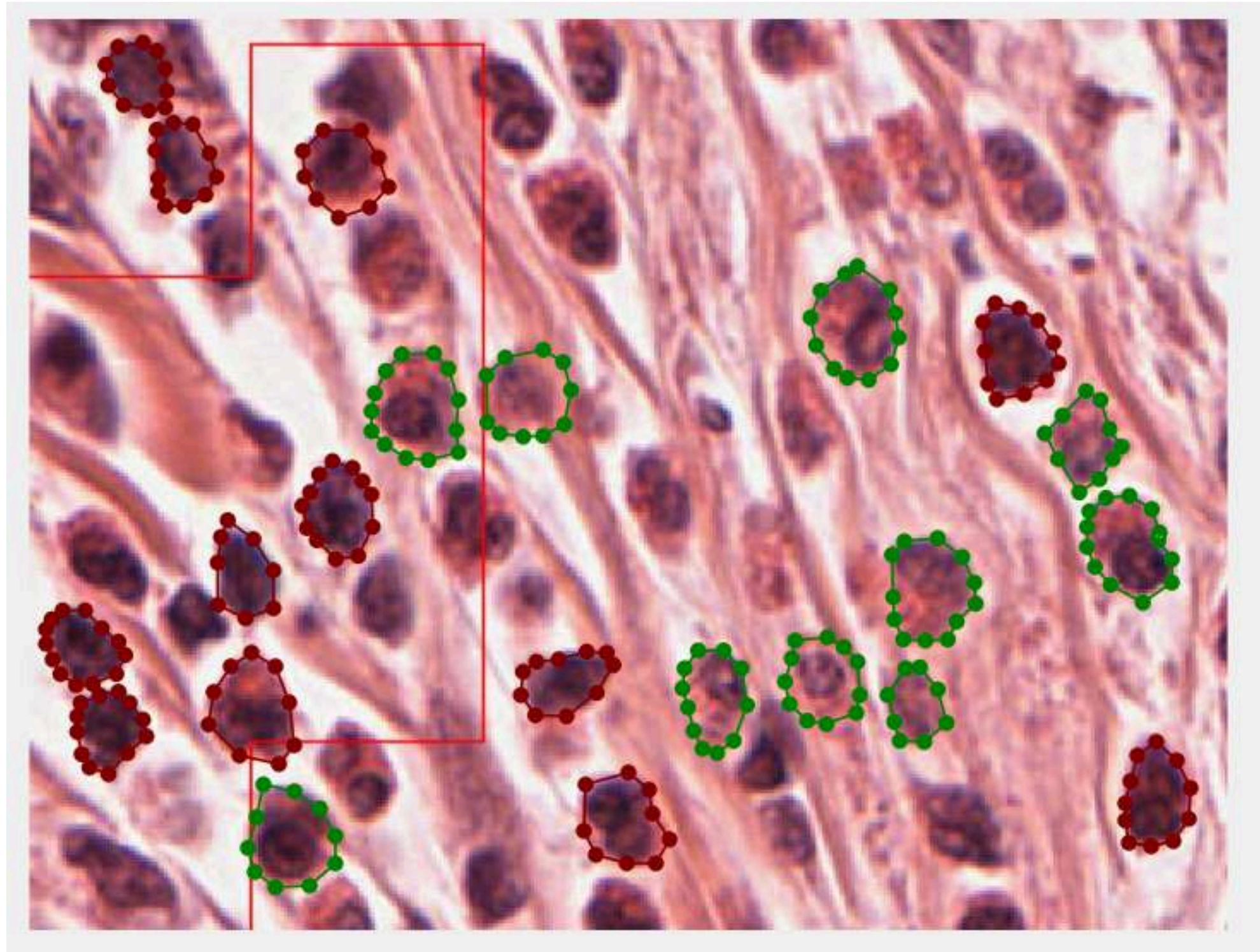
OBTENCION DE IMAGENES

- Utilizamos la imagen **P19-2717B.svs** y para tener las imágenes en buena resolución tomamos las capturas de pantalla con ayuda de File > Export snapshot > Current viewer content (SVG), para después en una página de internet convertirlas a JPG.



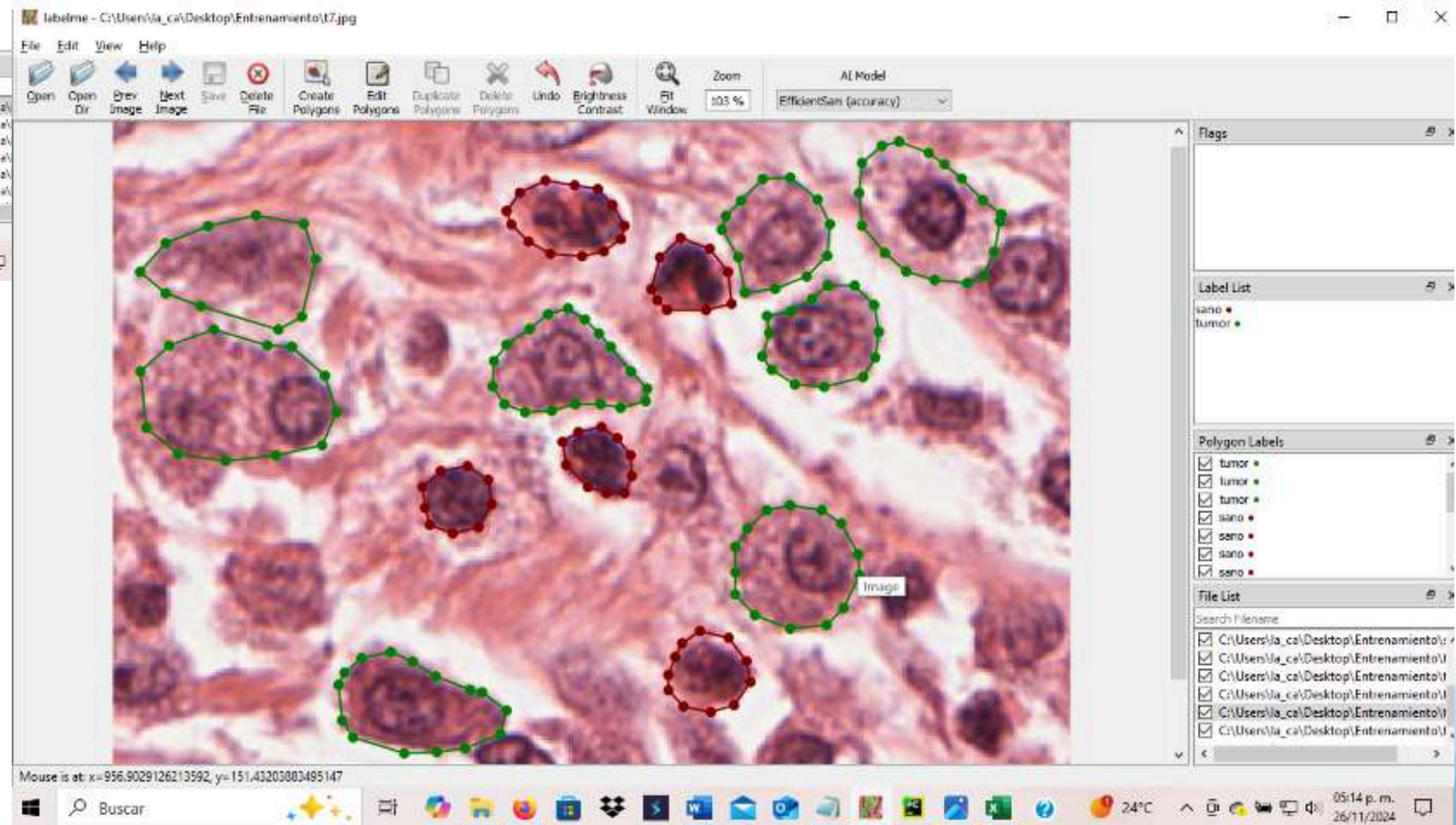
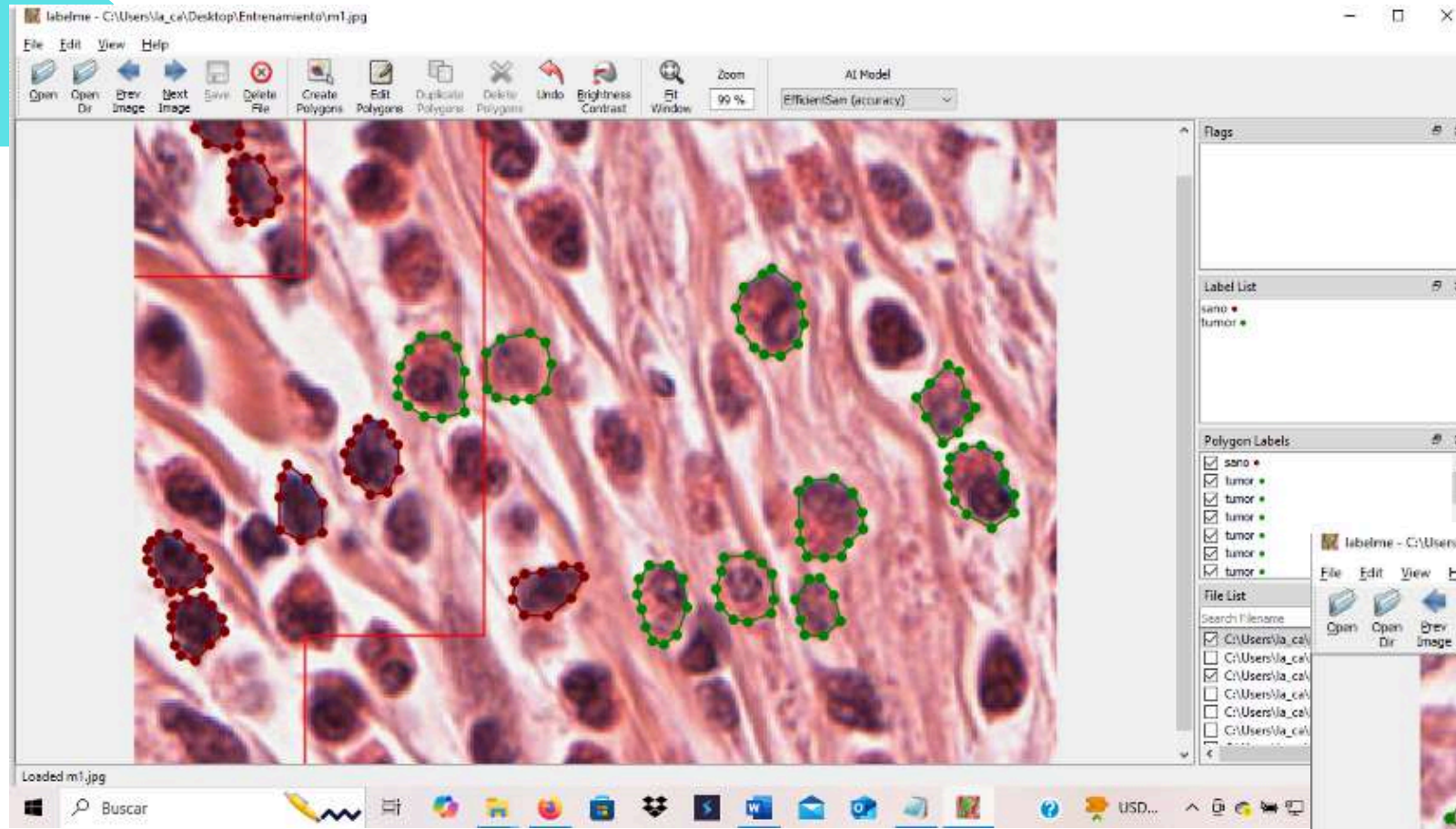
RECOLECCIÓN

Se utilizó la biblioteca labelme



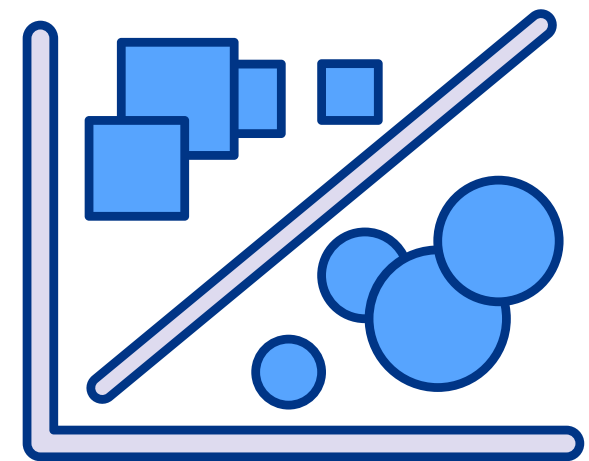
s2
s2
s5
s5
s6
s6
s10

OBTENCION DE IMAGENES



SVM COMO MODELO

- Teóricamente el mejor clasificador, debido a que encuentra un hiperplano que maximiza la separación entre clases, lo que puede mejorar la generalización.
- Es efectivo incluso si el número de características es mayor que el número de muestras, lo cual es común en problemas de imágenes.
- Tiene la ventaja de que por medio del Kernel transforma los datos de entrada a un espacio de características de mayor dimensión, lo que permite separar las clases de manera más efectiva.



EXTRACCIÓN DE CARACTERÍSTICAS



Contraste



Entropía



Área

Empezamos usando dos características para ir probando el desempeño del modelo; el contraste y la correlación. Esto nos dio métricas muy bajas, del 55%. Usamos todas las características disponibles en GLCM y algunas personalizadas como el área.

SET DE IMAGENES

Labelme nos devuelve un archivo json con la máscara que se le aplicará a la imagen para solo tomar en cuenta esa región con su respectiva etiqueta.

```
{
  "version": "5.5.0",
  "flags": {},
  "shapes": [
    {
      "label": "sano",
      "points": [
        [
          556.0,
          218.0
        ],
        [
          558.0,
          200.0
        ],
        [
          574.0,
          193.33333333333334
        ],
        [
          590.0,
          198.0
        ],
        [
          596.6666666666666

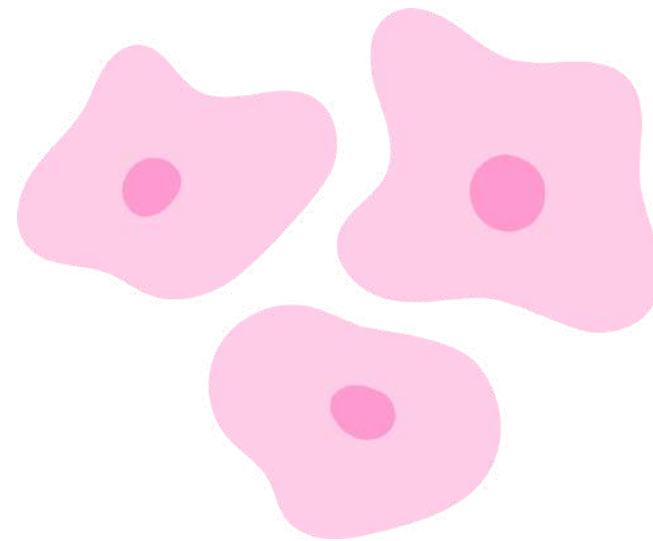
```

Así que hicimos más robusta la información aumentando primero el set de datos etiquetando más células, y después aumentando el número de características usando todas las disponibles de GLCM.

AÑADIENDO CARACTERÍSTICAS

Usamos características no definidas en GLCM como el **área de la célula**, así como la **media y desviación estándar** ya que creemos que con esto podemos medir un poco la granularidad que comenta el Dr; las células malas son menos “uniformes” que las tumorales que se pueden ver con texturizado más prominente.

Fue así como hicimos la elección de características que nos llevó a obtener métricas del 94%.



ENTRENAMIENTO Y PRUEBAS

	precision	recall	f1-score	support
sano	0.89	1.00	0.94	17
tumor	1.00	0.88	0.93	16
accuracy			0.94	33
macro avg	0.95	0.94	0.94	33
weighted avg	0.95	0.94	0.94	33
Precisión global: 0.94				

Con 200 datos de entrenamiento y 33 de prueba obtenemos los siguientes resultados. Aumentando las características y también el set de datos pasamos de 55% a 94%, pasando también por 60% y 82%.



DETECCIÓN

Semantic Segmentation



GRASS, CAT,
TREE, SKY

No objects, just pixels

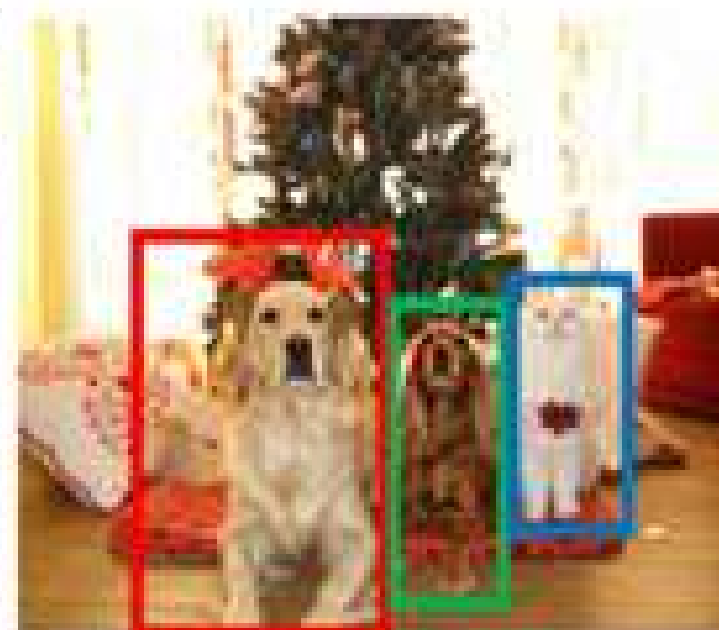
Classification + Localization



CAT

Single Object

Object Detection



DOG, DOG, CAT

Multiple Object

Instance Segmentation



DOG, DOG, CAT

La detección de objetos es una técnica de inteligencia artificial (IA) que se encarga de localizar y reconocer objetos en imágenes o videos.



DETECCIÓN

Por último quisimos implementar una técnica de detección para probar con una imagen que el algoritmo no conoce y ver cómo se desempeña.

En primera instancia intentamos seleccionar el contorno de todas las células, pero siempre había inconsistencias por lo que nos decidimos por detectar células y meterlas en una caja.

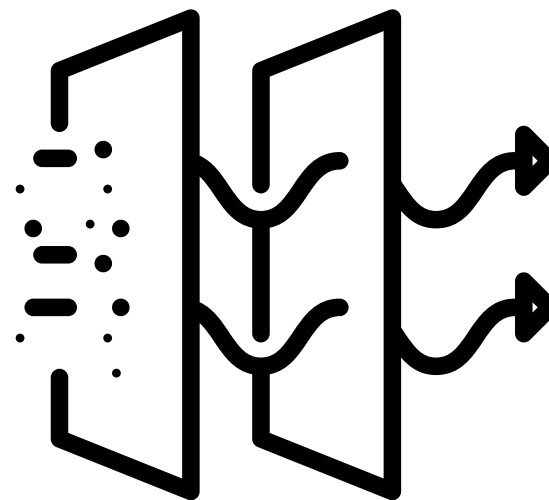
Una vez en la caja, se extraerán las características de las células detectadas y se meterán al algoritmo entrenado para ver cómo clasifica.



DETECCIÓN

Se hace un filtrado de los objetos detectados en el rango de 1000 a 12000 de área; para que no detecté cosas minúsculas que pueden corresponder únicamente a ruido.

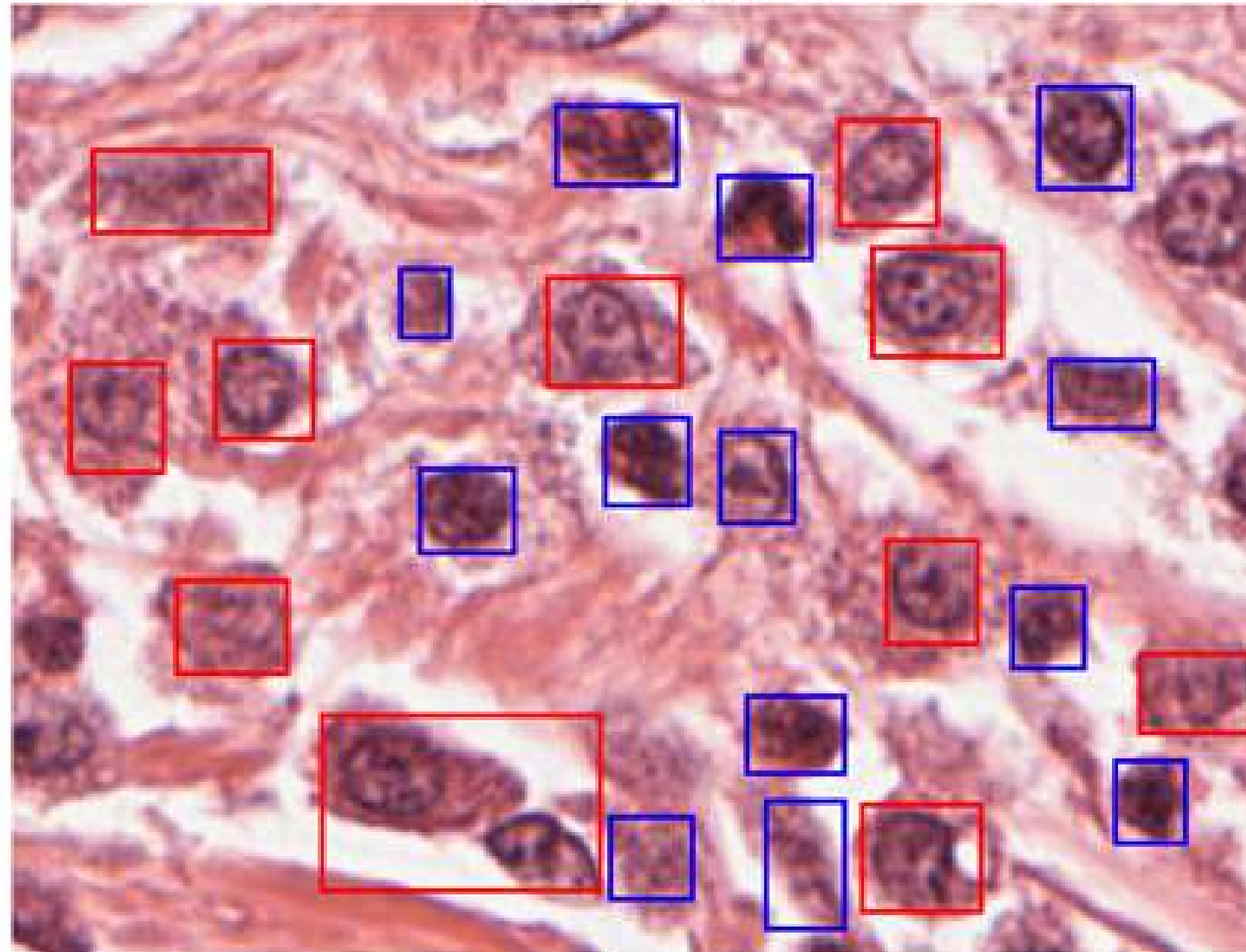
En un principio también filtramos por excentricidad pero resultó casi sin impacto.



RESULTADOS

- Podemos ver resultados prometedores a pesar de no tener un acierto del 100%. Estamos usando una imagen sin etiquetas completamente desconocida por el modelo.

Células detectadas: 24



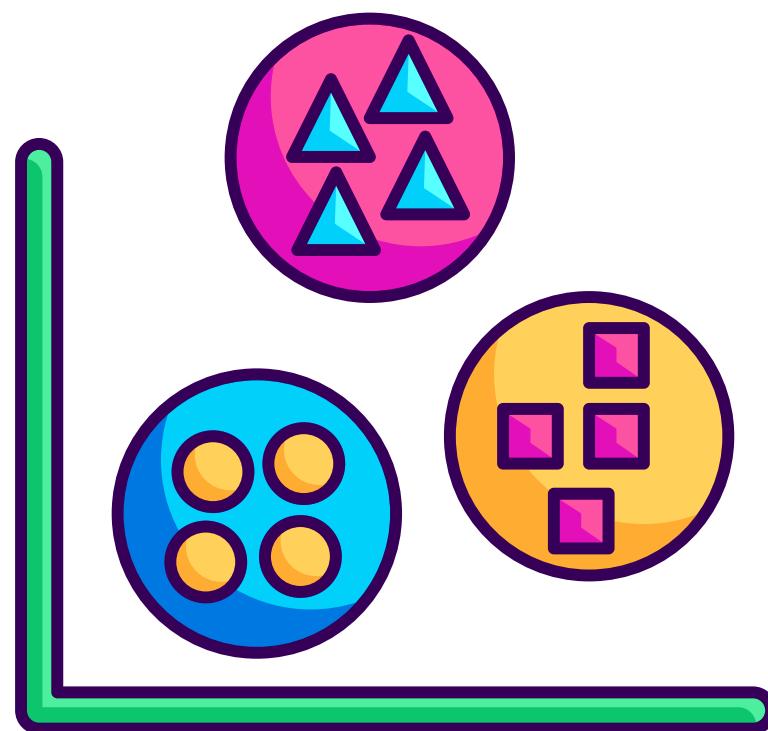
Rojo (mastocitos)
Azul (sana)

MEJORAS?

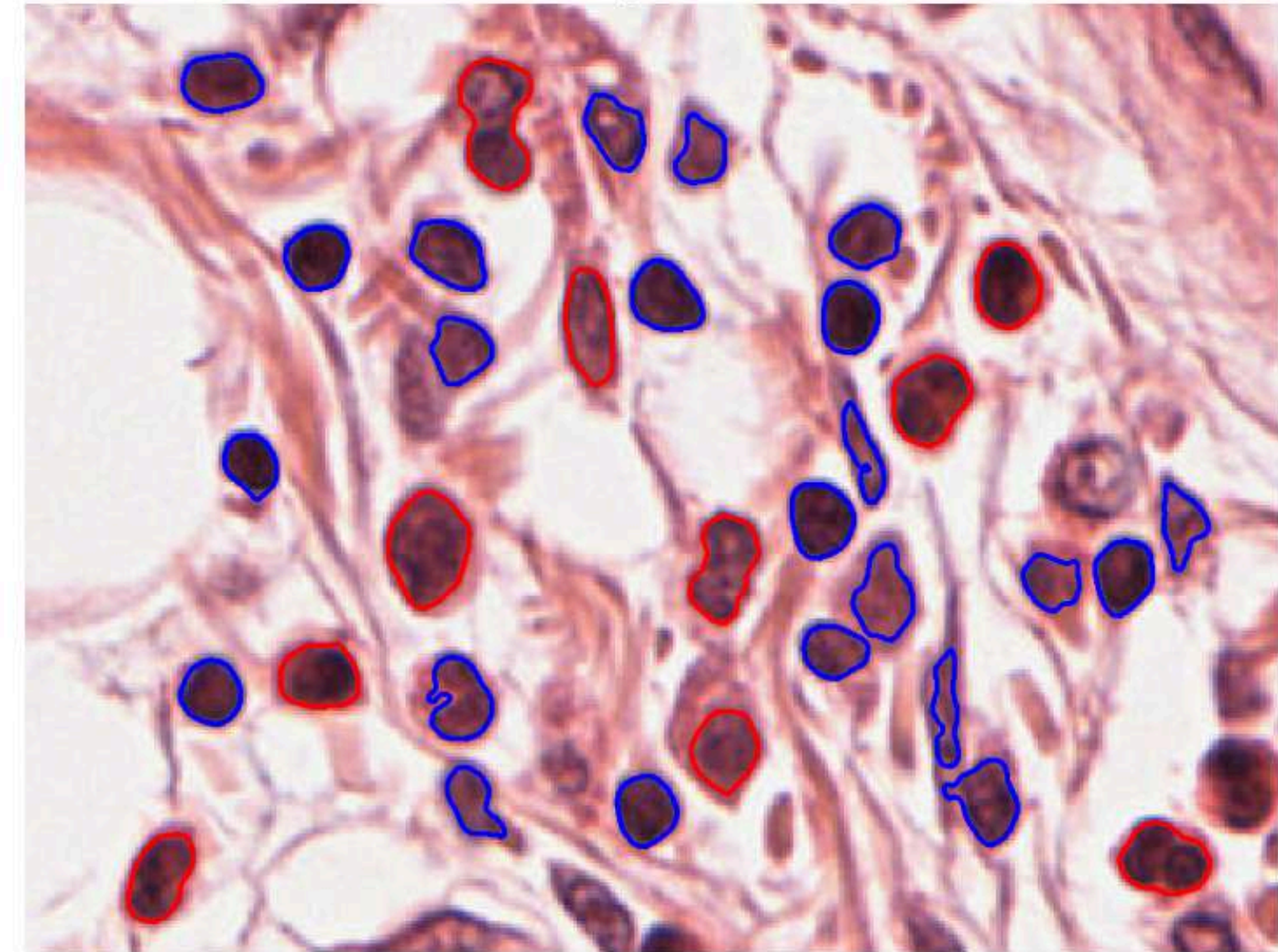
- Teniendo mayor conocimiento sobre las células es mas probable que no haya errores en nuestras etiquetas.
- Mas datos, promueven una mejor generalización hablando en áreas de machine learning.
- La detección es una tarea difícil y puede ser estudiada más a profundidad para detectar automáticamente las células en una imagen y pasarlas por el modelo entrenado; de esto dependen en gran medida los resultados que estamos obteniendo.

OTRAS PRUEBAS REALIZADAS

Probamos inicialmente la implementación de K-means pero no fue muy buena opción debido al gran parecido entre las dos clases. El modelo hacía predicciones como si las clases estuvieran totalmente balanceadas.



Clusters en la imagen de entrenamiento



RECONOCIMIENTO DE PATRONES

GRACIAS

3

- García López Erik
- Reyes Herrera Rogelio
- Serapio Hernández Alexis Arturo