

## “Software para conteo y clasificación automática de espermatozoides procesados mediante técnicas de HOS/Coomassie”

Córdoba, M.<sup>a</sup>; Giuliano, N.<sup>a</sup>; Miragaya, M.<sup>b</sup>; Plaza, J.<sup>b</sup>, Gerard, M.<sup>a</sup>

<sup>a</sup>. Instituto de Investigación en Señales, Sistemas e Inteligencia Computacional, sinc(i)-UNL/CONICET

<sup>b</sup>. Cátedra de Teriogenología, Facultad de Ciencias Veterinarias, Universidad de Buenos Aires.

mlcordoba26@gmail.com

### **Resumen**

La técnica combinada de HOS/Coomassie es utilizada para la evaluación de la calidad de espermatozoides de distintas especies. Este estudio es realizado por un operario entrenado y requiere del conteo manual de las células de cada muestra. Esta tarea demanda mucho tiempo, es repetitiva y se torna tediosa cuando la cantidad de muestras es elevada.

En este trabajo se propone el desarrollo de una herramienta de software basada en redes neuronales profundas para automatizar la tarea. La misma consiste de dos redes convolucionales concatenadas, ambas de la familia YOLO. Mientras que la primera realiza la segmentación y clasifica los espermatozoides individuales de acuerdo al test de HOS, la segunda red realiza la clasificación de acuerdo al test de Coomassie. Ambas fueron entrenadas y probadas con 633 y 15 imágenes, respectivamente. Los resultados experimentales muestran que la red HOS alcanza una tasa de clasificación de 87.7%, mientras que la red Coomassie logra una tasa de 91.8%, los cuales fueron los mejores resultados obtenidos respecto a pruebas realizadas junto con otras redes como Faster R-CNN, RetinaNet, EfficientDet y DSSD.

### **Abstract**

The combined HOS/Coomassie technique is used for evaluation of sperm quality in different species. This study is performed by a qualified operator and requires manual cells counting for each sample. This is a repetitive, time consuming, and tedious task, especially when there are a large number of samples.

In this work we propose the development of a software tool based on deep neural networks to automate the task. It consists of two concatenated convolutional networks, both of the YOLO family. While the first network performs the segmentation and classifies the individual sperm according to the HOS test, the second network performs the classification according to the Coomassie test. Both were trained and tested with 633 and 15 images, respectively. The experimental results show that the HOS network reaches a classification rate of 87.7%, while the Coomassie network achieves a rate of 91.8%, which were the best results obtained among other tests done with other networks such as Faster R-CNN, RetinaNet, EfficientDet and DSSD.

**Palabras clave:** Redes neuronales profundas, conteo de espermatozoides, HOS, Coomassie.

## INTRODUCCIÓN

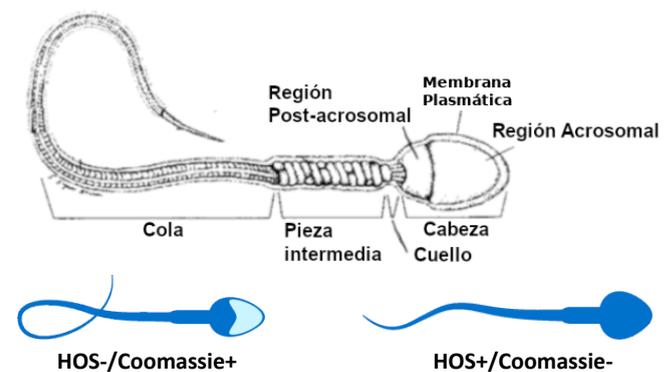
En la actualidad, los sistemas de análisis de esperma requieren de intervención humana para una correcta evaluación de la morfología de los espermatozoides [1], [2]. La Figura 1 *Arriba* esquematiza las regiones más importantes de la morfología de un espermatozoide sano. Para este análisis, se observan espermatozoides de diferentes muestras, y para cada uno se determina el resultado positivo o negativo para el Test de HOS y para la tinción de azul de Coomassie. El test de HOS evalúa la capacidad que posee la membrana plasmática para permitir el ingreso de agua a la célula. El mismo se realiza observando el engrosamiento y enrollamiento de la cola de los espermatozoides, y de esta forma determinar la funcionalidad de su membrana plasmática. Así son clasificados como *HOS positivos* si poseen algún grado de enrollamiento o engrosamiento de la cola, o *HOS negativos* en caso contrario (ver Figura 1 *Abajo*). El test de tinción de azul brillante de Coomassie (o test de Coomassie) se utiliza para la evaluación de la región acrosomal. La criopreservación del semen puede alterar los resultados de fertilidad esperados. Para evaluar dicha reacción, se utiliza el test de Coomassie, el cual se basa en observar dos patrones de coloración en la cabeza del espermatozoide. Como se observa en la Figura 1 *Abajo*, un espermatozoide puede clasificarse como *Coomassie positivo* (o Acrosoma positivo) si su región acrosomal tiene una coloración azul pareja en su totalidad. De otra forma, si sus regiones post-acrosomal y acrosomal poseen tonalidades de distinta intensidad, el espermatozoide se lo clasifica como *Coomassie negativo* (o Acrosoma negativo).

Desafortunadamente, analizar imágenes de espermatozoides demanda mucho tiempo, es repetitiva y se torna tediosa cuando la cantidad de muestras es elevada. Por este motivo, la automatización de cualquiera de las etapas del proceso podría contribuir a reducir los tiempos de procesamiento, estandarizar los procedimientos y reducir la variabilidad en los resultados.

El aprendizaje automático tiene una amplia gama de aplicaciones, incluyendo diagnósticos médicos, motores de búsqueda, análisis del

mercado de valores, clasificación de secuencias de ADN, reconocimiento del habla y del lenguaje escrito, juegos y robótica [3]. En particular, los modelos neuronales convolucionales son una incorporación reciente a este ecosistema, y han sido de gran utilidad para abordar problemas que involucran la visión por computadora, permitiendo la segmentación, clasificación y extracción de características de diferentes tipos de objetos presentes en imágenes. En particular, su uso en aplicaciones médicas ha proporcionado resultados sobresalientes [4].

En lo que respecta al análisis de espermatozoides, las redes neuronales convolucionales (o CNN, por sus siglas en inglés) han permitido automatizar muchas tareas de segmentación y clasificación gracias a su capacidad para extraer características de forma automática [5], [6]. Sin embargo, estos enfoques no han sido utilizados para abordar la clasificación automática de espermatozoides empleando el test de HOS/Coomassie. Teniendo en cuenta esto, en el presente trabajo se propone el desarrollo de una herramienta computacional que implemente un modelo neuronal capaz de automatizar la segmentación y clasificación de espermatozoides asegurando un conteo confiable.



**Figura 1: Arriba: Estructura de un espermatozoide<sup>1</sup>. Abajo: Ejemplos de clasificación de espermatozoides por la técnica HOS/Coomassie**

## PROPUESTA

La propuesta llevada adelante en este trabajo se basa en un algoritmo para detección de objetos en tiempo real ampliamente utilizado en

<sup>1</sup> Imagen provista por el Instituto de Investigación y Tecnología en Reproducción Animal: <http://www.fvet.uba.ar/?q=initra#marco>

la literatura. Debido a esto, primero se presenta la arquitectura de este modelo neuronal y se describen sus componentes. Luego se presentan los dos modelos propuestos bajo esta arquitectura para la clasificación del test HOS/Coomassie: i) modelo de *red única*, que realiza la segmentación y clasificación de espermatozoides utilizando una sola red convolucional entrenada; ii) modelo de *red combinada*, que emplea redes convolucionales con igual arquitectura que se acoplan en cadena.

### Arquitectura YOLO

YOLO corresponde a la abreviatura del término 'You Only Look Once' (o en español, 'solo se mira una vez'). Es un modelo neuronal basado en CNN capaz de segmentar y clasificar varios objetos presentes en una imagen en una única ejecución del algoritmo [7]. La detección de objetos en YOLO se resuelve mediante un problema de regresión. Además, este modelo es capaz de proporcionar la probabilidad de pertenencia de un objeto a las diferentes clases consideradas en el problema.

La arquitectura general de este modelo, cuenta con 24 capas convolucionales, seguidas de 2 capas completamente conectadas. Alterna capas convolucionales de tamaño 1x1 para combinar información de los canales de la imagen, y 3x3 para reducir la cantidad de datos respecto a las capas anteriores. Este modelo además incorpora tres elementos adicionales en su arquitectura: *Bloques residuales*, *Regresión del cuadro delimitador* e *Intersección sobre unión (IoU)*.

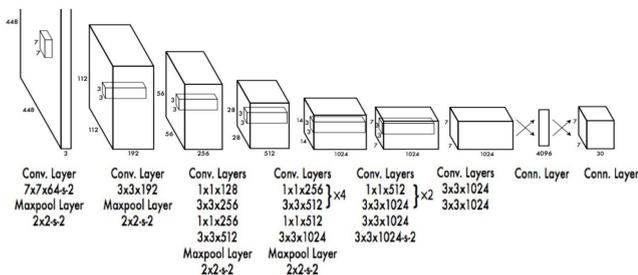


Figura 3: Arquitectura YOLO<sup>2</sup>.

La Figura 3 esquematiza los elementos más importantes de esta arquitectura. Las capas

<sup>2</sup>[https://www.researchgate.net/figure/YOLOv3-object-detection-architecture\\_fig5\\_340903596](https://www.researchgate.net/figure/YOLOv3-object-detection-architecture_fig5_340903596)

convolucionales iniciales extraen características de la imagen y las capas finales predicen el grado de certeza de la segmentación y coordenadas del objeto [7].

Respecto a los elementos adicionales esta red incorpora un Bloque residual para solucionar errores de precisión por saturación, los cuales son típicos en las redes profundas. La diferencia arquitectónica entre el bloque de convolución normal y el bloque residual es la adición de la conexión de salto. La conexión de salto lleva la entrada a las capas más profundas para de esta forma mejorar el entrenamiento y la propagación del gradiente<sup>3</sup>.

Por otro lado, la técnica de *Regresión del cuadro delimitador* utiliza cuadros (o cajas) con contornos que resaltan los objetos de interés en la imagen. Cada cuadro delimitador contiene información sobre el ancho, alto, coordenadas del centro del cuadro y la clase que representa.

Por último la medida IoU describe cómo se superponen estos cuadros. YOLO usa esta técnica para proporcionar un cuadro delimitador de salida que rodea los objetos de la forma más precisa posible. La Figura 5 muestra un ejemplo de aplicación de esta medida. Se pueden observar dos cuadros delimitadores, uno en rojo y el otro en azul. El cuadro azul es el cuadro predicho, mientras que el cuadro rojo es el cuadro real. YOLO asegura que los dos cuadros delimitadores sean iguales.

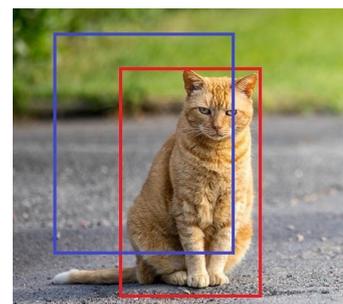


Figura 5: Imagen con dos cajas superpuestas<sup>4</sup>

Cada celda de la cuadrícula es responsable de predecir los cuadros delimitadores y sus puntuaciones de confianza. Por ejemplo, el IoU es igual a 1 si el cuadro delimitador previsto es el

<sup>3</sup> <https://www.inverseai.com/blog/yolov3-object-detection/>

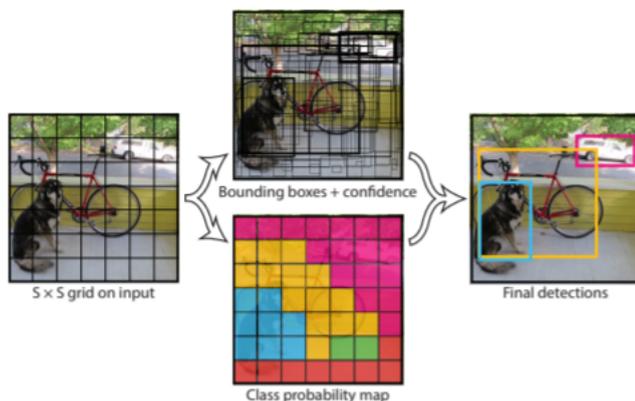
<sup>4</sup> <https://medium.com/koderunners/intersection-over-union-516a3950269c>

mismo que el cuadro real. Este mecanismo elimina los cuadros delimitadores que no son iguales al cuadro real.

La Figura 6 muestra cómo YOLO procesa cada imagen para producir los resultados de detección finales. Primero, la imagen se divide en una grilla de celdas. Cada una de estas pronostica los cuadros delimitadores y proporciona sus puntuaciones de confianza. Las celdas predicen las probabilidades de clase para establecer la clase de cada objeto.

La intersección sobre la unión asegura que los cuadros delimitadores predichos sean iguales a los cuadros reales de los objetos. Este fenómeno elimina los cuadros delimitadores innecesarios que no cumplen con las características de los objetos (como altura y ancho). La detección final consistirá en cuadros delimitadores únicos que se ajustan perfectamente a los objetos.

Finalmente cabe destacar que la función de costo utilizada por este modelo durante el entrenamiento es el gradiente descendente. El cual es un método de optimización numérica para estimar los mejores coeficientes, consiste en calcular el gradiente de la función de error del modelo y modificar los parámetros del mismo en el sentido opuesto, de forma tal de disminuir el error y mejorar el desempeño y cada interacción.



**Figura 6: Técnicas combinadas de YOLO para la detección y clasificación de objetos<sup>5</sup>**

## Modelos propuestos

En este trabajo se proponen dos modelos neuronales basados en la arquitectura YOLO.

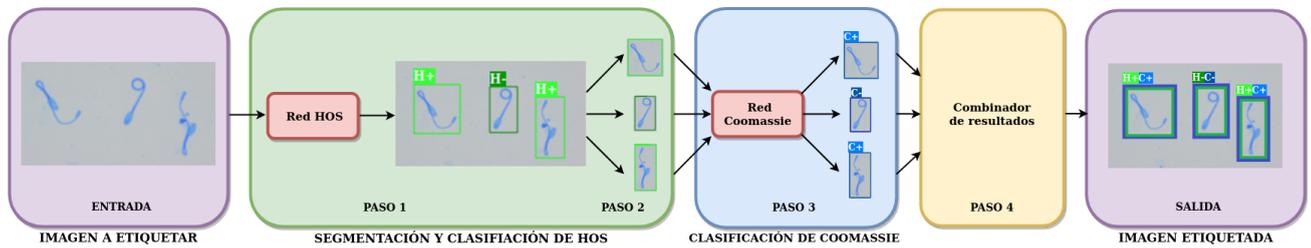
El primer modelo, denominado *red única*, consiste en una única red YOLO que es entrenada para realizar la segmentación y clasificación de ambos tests en un solo paso.

El segundo modelo, denominado *red combinada*, consiste en dos redes YOLO que son entrenadas en forma independiente, tal como se muestra en la Figura 7. La idea detrás de este enfoque es analizar si dividiendo la tarea de clasificación del test combinado mejora el desempeño de las redes. Cada imagen es procesada por la primera red, denominada *red HOS*, que se encarga de realizar la segmentación de los espermatozoides y la clasificación de acuerdo a la prueba de HOS. La imagen a la salida de esta red es recortada para producir parches que contienen cada uno de los espermatozoides segmentados (ver Paso 2 en la Figura 4). De estos recortes se descartan las imágenes que no sean clasificables, y las restantes son alimentadas a la segunda red, denominada red Coomassie. Está ahora clasifica cada espermatozoide como Coomassie positivo o negativo, de acuerdo a la coloración de la cabeza. Similar a la primera red, se tienen como salida las sub-imágenes clasificadas y etiquetadas, como se aprecia en el recuadro azul. Seguidamente, se toman las etiquetas generadas por ambas arquitecturas y se construye la etiqueta final que consta de los resultados de ambas clasificaciones, correspondientes a las etiquetas originales.

## RESULTADOS Y DISCUSIÓN

En esta sección se presentan los resultados experimentales obtenidos en las diferentes pruebas. Primero se realiza una comparación del desempeño para diferentes modelos neuronales utilizados típicamente en tareas de segmentación y clasificación. Luego se compara el desempeño entre las dos arquitecturas propuestas en este trabajo. Finalmente, se realiza la optimización de hiperparámetros para la red combinada y se analizan las mejoras de desempeño. Todos los experimentos se llevaron a cabo empleando 50 épocas de entrenamiento.

<sup>5</sup>[https://www.guidetomlandai-com.translate.google/assets/img/computer\\_vision/YOLO.PNG?\\_x\\_tr\\_sl=auto&\\_x\\_tr\\_tl=es&\\_x\\_tr\\_hl=es-419&\\_x\\_tr\\_pto=nui](https://www.guidetomlandai-com.translate.google/assets/img/computer_vision/YOLO.PNG?_x_tr_sl=auto&_x_tr_tl=es&_x_tr_hl=es-419&_x_tr_pto=nui)



**Figura 7: Proceso de segmentación, clasificación y obtención de etiquetas de la red combinada.**

### Dataset y medidas de evaluación

Para evaluar estos desempeños, se utiliza la medida Mean Average Precision (mAP) que se conforma como la media de todas precisiones medias de las clases, y se calcula como:

$$mAP = \frac{1}{N} \sum_{i=1}^N AP_i \quad (1)$$

donde AP es la Precisión Media (o en inglés, Average Precision). Así, si consideramos 5 imágenes correspondiente a 5 clases diferentes y se clasifican correctamente 4, el mAP será de 0.8.

El dataset empleado en este trabajo fue construido a partir de un conjunto de 633 imágenes provistas por el Instituto de Investigación y Tecnología en Reproducción Animal (o INITRA). Estas fueron obtenidas utilizando un microscopio de contraste de fase. Adicionalmente se necesitó que dichas imágenes estén etiquetadas tanto para la validación como para el proceso de entrenamiento.

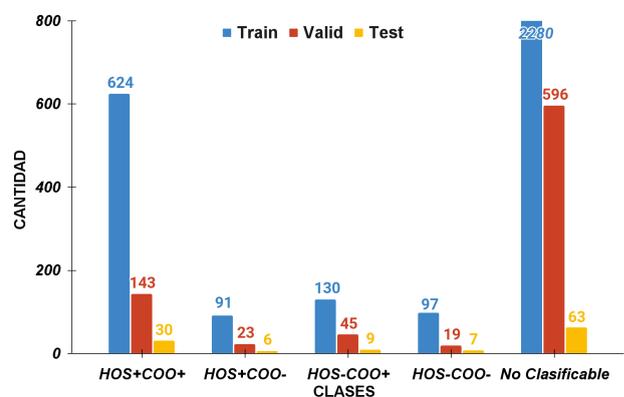
El etiquetado se realizó dividiendo los espermatozoides en cinco clases, lo que permite la clasificación combinada de los resultados para las pruebas HOS y Coomassie. Así, las 5 clases quedaron determinadas como H+C+, H+C-, H-C+, H-C-, NC. Donde *H* y *C* hacen referencia al test de HOS y de Coomassie respectivamente, mientras que *NC* hace referencia a los casos *No Clasificables*. Por ejemplo, casos cuando se tienen varios espermatozoides solapados entre sí, o no se ven enteros debido a que aparecen en el borde de la imagen.

A partir del dataset se crearon tres grupos de imágenes: uno de entrenamiento, otro de validación y otro de test con 391, 100 y 142

imágenes respectivamente. La Figura 8 esquematiza la distribución de clases para cada grupo. Se observa en ellos un gran desbalance, por lo que se realizaron procesos de balanceo antes de efectuar los experimentos.

### Comparación de desempeño entre diferentes arquitecturas de redes convolucionales

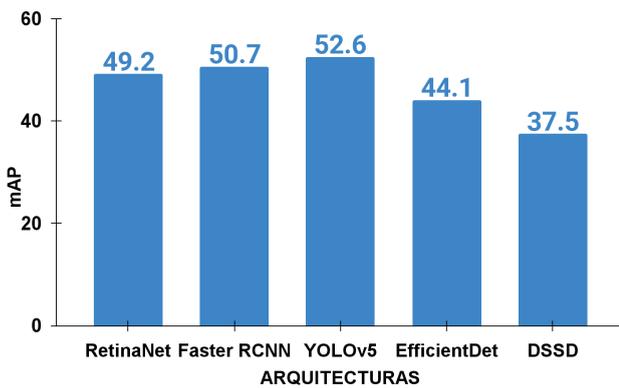
Existen diferentes arquitecturas capaces de segmentar y clasificar imágenes simultáneamente. Para determinar la más adecuada para este problema se realizó una comparación entre 5 arquitecturas ampliamente utilizadas en la literatura. En particular, para el caso de YOLO se empleó la implementación YOLOv5 proporcionada por Ultralytics<sup>6</sup>, que consiste en una versión mejorada del modelo YOLOv3. Todos los modelos fueron entrenados y evaluados empleando los mismos conjuntos de imágenes. Las Figuras 9 y 10 presentan los resultados comparativos entre estas arquitecturas.



**Figura 8: Distribución de clases en el dataset.**

<sup>6</sup> <https://github.com/ultralytics/yolov5>

Como puede apreciarse, YOLO presenta el mejor desempeño frente a las arquitecturas consideradas para el enfoque de red única. Por su parte, RetinaNet [8] y Faster RCNN [9] presentan resultados similares a YOLO, aunque un poco por debajo del mismo. En contrapartida, EfficientDet [10] y DSSD [11] presentan los desempeños más pobres, siendo el último modelo el que produce los peores resultados.



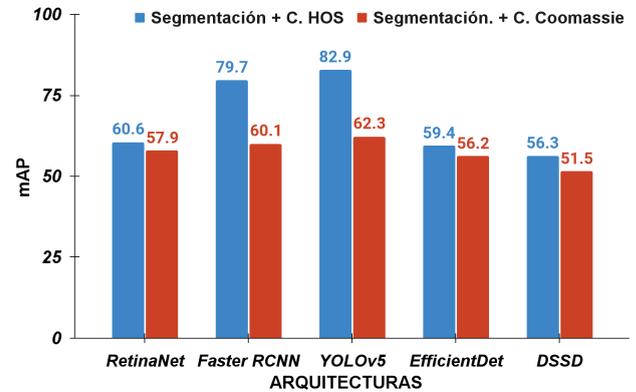
**Figura 9: Comparación de desempeños utilizando una única red para la segmentación y clasificación de ambos tests.**

Con la finalidad de mejorar los resultados anteriores, se propuso un enfoque alternativo en el cual se realice la segmentación y clasificación de HOS por un lado, y la segmentación y clasificación de Coomassie por otro. Para esto, se entrena nuevamente cada red, una vez para cada proceso de segmentación y clasificación. Los resultados obtenidos para los valores de mAP se presentan en la Figura 10. Se puede observar claramente que al separar las pruebas, los resultados mejoran notablemente. Cabe destacar que todos los modelos emplearon aproximadamente 2 horas para completar el entrenamiento.

Como se puede ver en la Figura 9, el proceso de segmentación y clasificación en un paso no se logran resultados satisfactorios, pero aún así, las pruebas fueron útiles para idear y desarrollar el segundo enfoque. A pesar de esto se puede ver que YOLOv5 es la arquitectura que presenta mejores resultados.

Como se observa en la Figura 10 las arquitecturas pueden segmentar correctamente si se realiza la segmentación y clasificación de

HOS por un lado, y la segmentación y clasificación de Coomassie por otro.



**Figura 10: Comparación de desempeños utilizando una red para la segmentación y clasificación de HOS y otra para la segmentación y clasificación de Coomassie.**

Lo visto anteriormente sugiere que utilizar esta red y dividir el proceso de clasificación del test combinado en dos tareas es el camino más adecuado para obtener implementaciones con mayor desempeño.

### Optimización de hiperparámetros

Se buscó optimizar el desempeño de la red combinada. Para esto se realizaron pruebas asignando diferentes valores al umbral de confianza, el cual determina el mínimo grado de certeza que debe tener una detección para que sea considerada en el análisis, y al coeficiente de IoU.

Estas pruebas fueron realizadas con los datos de entrenamiento, tanto para la red HOS como para la red Coomassie. La asignación de valores se llevó a cabo considerando todas las combinaciones posibles para estos parámetros, los cuales pueden ser optimizados para cada red. Los resultados se muestran en la Tabla 1 y 2 para la red HOS y red Coomassie, respectivamente. En ellas se muestran, a través de una matriz de colores, los valores de mAP obtenidos por cada posible combinación de valores de los parámetros mencionados.

La Tabla 1 muestra que los mejores valores de precisión para la red HOS se obtienen para el umbral de confianza cercano a 0 y el coeficiente de IoU entre 0.2 y 0.8. Adicionalmente puede observarse que la precisión decrece de forma

progresiva y constante a medida que se aumenta el valor del umbral de confianza.

Por otra parte, la Tabla 2 muestra que para la red Coomassie se presentan cambios bruscos de precisión en los valores máximos de estos parámetros, variando muy poco en el medio.

**Tabla 1: Resultados de pruebas de optimización red HOS.**

		Confianza										
		0.0	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	1.0
IoU	0.0	77.3	75.1	74.2	73.3	72	70.5	67.6	62	45.2	0.4	0
	0.1	84.7	82.1	80.1	79.9	78.5	76.9	73.6	67.1	47.9	0.4	0
	0.2	90.1	87.6	80.9	85.5	83.8	82	78.1	70.9	49.8	0.4	0
	0.3	93.1	90.6	89.5	88.3	86.5	84.6	80.5	72.6	50.7	0.4	0
	0.4	94.4	92.2	90.9	89.7	87.9	85.9	81.7	73.6	51.1	0.4	0
	0.5	95	93.1	91.9	90.6	88.8	86.7	82.4	74.1	51.3	0.4	0
	0.6	95.2	93.4	92.3	91	89.2	87	82.6	74.2	51.3	0.4	0
	0.7	95	93.4	92.3	91.1	89.3	87	82.7	74.3	51.3	0.4	0
	0.8	94.2	93	92	90.9	89.2	87.1	82.7	74.3	51.3	0.4	0
	0.9	87.6	87	86.6	86	85	83.5	80.1	72.7	50.8	0.4	0
	1.0	23.4	23.3	23.3	23.2	23	22.8	22.3	21.1	16.8	0.2	0

La Tabla 3 presenta la comparación de los valores de mAP obtenidos para la red única y para la red combinada antes y después de ser optimizada. Los valores de mAP individual y

**Tabla 2: Resultados de pruebas de optimización red Coomassie.**

		Confianza										
		0.0	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	1.0
IoU	0.0	98.7	98.8	98.1	97.8	97.2	96.2	95.3	94.6	90.4	68.7	0
	0.1	98.7	98.8	98.1	97.8	97.2	96.2	95.3	94.6	90.4	68.7	0
	0.2	98.7	98.8	98.1	97.8	97.2	96.2	95.3	94.6	90.4	68.7	0
	0.3	98.7	98.8	98.1	97.8	97.2	96.2	95.3	94.6	90.4	68.7	0
	0.4	98.7	98.8	98.1	97.8	97.2	96.2	95.3	94.6	90.4	68.7	0
	0.5	98.7	98.8	98.1	97.8	97.2	96.2	95.3	94.6	90.4	68.7	0
	0.6	98.7	98.8	98.1	97.8	97.2	96.2	95.3	94.6	90.4	68.7	0
	0.7	98.7	98.8	98.1	97.8	97.2	96.2	95.3	94.6	90.4	68.7	0
	0.8	98.6	98.8	98.1	97.8	97.2	96.2	95.3	94.6	90.4	68.7	0
	0.9	97.6	97.7	97.1	96.9	96.3	95.5	94.5	93.9	89.8	68.5	0
	1.0	39.8	39.8	39.8	39.8	39.7	39.6	39.6	39.5	39.1	35.1	0

combinado presentados en la tabla corresponden a los casos donde se analiza el desempeño de las redes por separado y en conjunto considerando la clasificación en 5 clases (ver etiquetas del dataset). Para la red única y la red combinada se utilizaron valores fijos de 0.25 y 0.45 para el umbral de confianza e IoU,

respectivamente, y para la red optimizada se usaron los valores más altos que se muestran en las Tablas 1 y 2. Cabe mencionar que el tiempo de entrenamiento fue similar en todos los casos rondando nuevamente las 2 horas.

**Tabla 3: Resultados de las pruebas realizadas con el dataset test bajo las 5 clases establecidas.**

		Red Única	Red Combinada		Red Combinada Optimizada	
			Red HOS	Red Coomassie	Red HOS	Red Coomassie
mAP	Individual	54.7	81.6	88.5	87.7	91.8
	Combinado		72.2		80.5	

Puede observarse que el mAP combinado para la red combinada y la red optimizada, aumenta 17.5 y 25.8 puntos respecto de la red única, respectivamente. A su vez, la optimización de la red combinada proporciona una mejora notable respecto del modelo de red combinada sin optimizar. Claramente esto resalta la importancia de estos hiperparámetros sobre el desempeño del modelo propuesto.

### Interfaz de usuario

Para facilitar el uso de la herramienta, se encuentra en versión *beta* una interfaz de usuario mediante la librería *Streamlit*. La Figura 6 presenta una captura de dicha interfaz.



**Figura 6: Interfaz de usuario del software.**

Esta interfaz es simple, minimalista y contiene los elementos indispensables para realizar la tarea. A su vez se divide en cinco secciones: un botón para cargar las imágenes a procesar; barra deslizante para modificar el umbral de confianza; un botón para iniciar el proceso de detección y clasificación; un gráfico de barras que muestra los resultados de la clasificación; un botón para descargar las imágenes procesadas. Una versión de prueba se encuentra disponible en

<https://hare.streamlit.io/solutionsai/detecthc/main.py>.

## CONCLUSIONES

En este trabajo se desarrolló un software para conteo y clasificación automática de espermatozoides procesados mediante técnicas de HOS/Coomassie. Para llevar adelante las pruebas se construyó una base de datos con 633 imágenes, y a partir del mismo se crearon tres grupos de imágenes: uno de entrenamiento, otro de validación y otro de test.

Se evaluaron 6 modelos neuronales reportados en la literatura adecuados para segmentación y clasificación. Se encontró que el modelo YOLO presenta el mejor desempeño. Este modelo fue empleado para evaluar dos enfoques: red única y red combinada. Mientras que el primero resuelve el problema de clasificación de ambos test al mismo tiempo, el segundo utiliza una red independiente para analizar cada test. Se observó que la descomposición del problema aumenta el desempeño del modelo neuronal, alcanzando mejores tasas de clasificación.

Adicionalmente se desarrolló una simple interfaz de usuario empleando una librería de uso libre, que permite un fácil uso de la implementación para los investigadores del INITRA u otros posibles usuarios del sistema.

## REFERENCIAS

- [1] Auger, J.; Jouannet, P.; Eustache, F. (2016) Another look at human sperm morphology. Hum Reprod
- [2] Van den Hoven, L.; Hendriks, J.; Verbeet, J.; Westphal, J.; M Wetzels, A. (2014) Status of

sperm morphology assessment: an evaluation of methodology and clinical value. Elsevier.

- [3] Mitchell, T. (1997). *Machine Learning*, McGraw Hill.
- [4] Garcia-Garcia, A.; Orts-Escolano, S.; Oprea, S.O.; Villena-Martinez, V. (2017). A Review on Deep Learning Techniques Applied to Semantic Segmentation. arXiv:1704.06857v1 [cs.CV].
- [5] Shaker, F.; Amirhassan Monadjemi, S.; Reza Naghsh-Nilchi, A.; (2016). Automatic detection and segmentation of sperm head acrosome and nucleus in microscopic images of human semen smears. Elsevier.
- [6] Movahed, R.; Mohammadi, E.; Orooji, M. (2017). Automatic segmentation of Sperm's parts in microscopic images of human semen smears using concatenated learning approaches. Computers in Biology and Medicine, volumen 109, 242-253.
- [7] Redmon, J.; Divvala, S.; Girshick, R.; Farhadi, A. (2016) You Only Look Once: Unified, Real-Time Object Detection. Facebook AI Research. arXiv:1506.02640v5 [cs.CV]
- [8] Lin, T.; Goyal, P.; Girshick, R.; He, K.; Dollár, P. (2018) Focal Loss for Dense Object Detection. arXiv:1708.02002v2 [cs.CV].
- [9] Ren, S.; He, K.; Girshick, R.; Sun, J. (2016) Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks. arXiv:1506.01497v3 [cs.CV].
- [10] Tan, M.; Le, V. (2020) EfficientNet: Rethinking Model Scaling for Convolutional Neural Networks. arXiv:1905.11946v5 [cs.LG].
- [11] Fu, C.; Liu, W.; Ranga, A.; Tyagi, A.; Berg, A. (2017) DSSD: Deconvolutional Single Shot Detector. arXiv:1701.06659 [cs.CV].

# SOFTWARE PARA CONTEO Y CLASIFICACIÓN AUTOMÁTICA DE ESPERMATOZOIDES PROCESADOS MEDIANTE TÉCNICAS COMBINADAS DE HOS/COOMASSIE

N. Giuliano, M. Martín y M.F. Gerard

Instituto de Investigación en Señales, Sistemas e Inteligencia Computacional (sinc(i)), FICH-UNL/CONICET, Argentina

## Introducción

Para evaluar la calidad de espermatozoides de distintas especies se utiliza frecuentemente la técnica combinada de HOS/Coomassie (HOS/CB). Las posibles configuraciones para esta técnica se ilustran en la Figura 1.

El análisis de la técnica se realiza manualmente, contando los espermatozoides que presentan un aspecto positivo o negativo en cada caso.

Esta tarea es repetitiva y tediosa cuando la cantidad de muestras es elevada.

Las soluciones basadas en visión por computadora basadas en aprendizaje profundo pueden ser una alternativa a esta solución a este proceso.

En este trabajo se desarrolla una herramienta de software basada en redes neuronales profundas para automatizar la tarea.

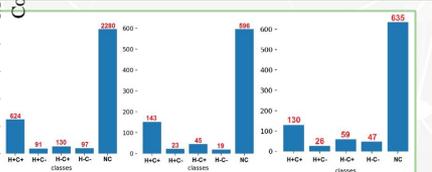


**HOS Positivo:** colas enrolladas o engrosadas.  
**HOS Negativo:** colas sin enrollar y sin engrosamiento.

**Coomassie Positivo:** cabezas con coloración azul homogénea.  
**Coomassie Negativo:** cabezas con dos tonalidades de azul diferente.

Figura 1. Posibles resultados del test combinado.

## Dataset



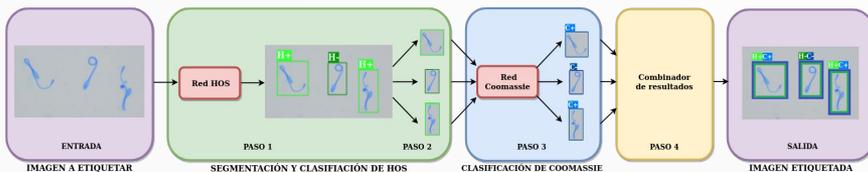
Se dispone de un dataset de 633 imágenes etiquetadas en 5 clases (H+C+, H-C+, H+C-, H-C-, No Clasificable) que se dividieron a su vez en 3 datasets, entrenamiento, validación y test. Se observa en ellos un gran desbalance, por lo que se realizaron procesos de balanceo y se evaluaron diferentes modelos para mitigarlo.

## Interfaz

Para facilitar el uso de la herramienta, se encuentra en desarrollo una interfaz de usuario, mediante la librería Streamlit. Esta permite cargar imágenes, modificar algunos parámetros del proceso de detección y generar un gráfico con los resultados generales de las clasificaciones. También permite descargar los resultados. La versión preliminar de la interfaz puede accederse mediante el siguiente código QR.

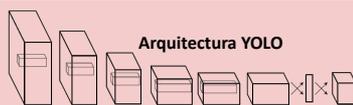


## Algoritmo Propuesto



**Redes HOS y Coomassie:** Modelos neuronales convolucionales pertenecientes a la familia YOLO

## Arquitectura YOLO



**ENTRADA:** Imagen original con espermatozoides a segmentar y etiquetar

**PASO 1-** Proceso de segmentación y clasificación de HOS de espermatozoides sobre la imagen de **ENTRADA**  
 En este primer paso se busca segmentar cada uno de los espermatozoides presentes en la imagen y etiquetarlos bajo la técnica de HOS

**SALIDA 1:** Imagen segmentada y etiquetada por la red HOS

**PASO 2-** Recorte de los espermatozoides y generación de imágenes individuales  
 Se genera una sub-imagen con la imagen segmentada de cada espermatozoide

**PARACADA** etiqueta  $\alpha$  de **SALIDA 1**  
 generar sub-imagen cuyas dimensiones coinciden con las de la etiqueta  $\alpha$

**SALIDA 2:**  $n$  sub-imágenes etiquetadas, donde  $n$  denota la cantidad de espermatozoides segmentados en **PASO 1**

**PASO 3-** Clasificación de Coomassie de espermatozoides sobre las sub-imágenes generadas  
 En esta etapa se etiquetan cada uno de los recortes generados anteriormente bajo la técnica de Coomassie

**PARACADA** sub-imagen  $k$  de **SALIDA 2**  
 Si clase de  $k$  es 'NC'  
 remover sub-imagen  $k$  y su respectiva etiqueta  
**SINO**  
 clasificar y etiquetar sub-imagen  $k$  bajo el test de Coomassie

**SALIDA 3:**  $m$  sub-imágenes etiquetadas bajo test de Coomassie, donde  $m \leq n$

**PASO 4-** Combinación de resultados  
 Aquí se combina la clasificación de cada una de las redes para cada espermatozoide segmentado

**PARACADA** etiqueta  $\alpha$  de **SALIDA 1**  
 Si etiqueta  $\alpha$  no es 'NC'  
 se le concatena a  $\alpha$  la etiqueta resultante de la clasificación de la sub-imagen correspondiente en el **PASO 3**

**SALIDA:** Imagen con espermatozoides segmentados y clasificados bajo la técnica combinada HOS/Coomassie

## Resultados

- Se entrenó el modelo utilizando dos esquemas:  
**1- Red Única:** Se tiene una sola red entrenada para segmentar y clasificar las imágenes bajo el test combinado de HOS/Coomassie.  
**2- Red Combinada:** Se tiene una red entrenada para segmentar y realizar la clasificación de imágenes de acuerdo al test de HOS. Luego se procesa la salida con una segunda red, encargada de clasificar de acuerdo al test de Coomassie.

- Para determinar el desempeño de los esquemas se utiliza la métrica Mean Average Precision (mAP), la cual se conforma como la media de todas precisiones medias de las clases.

- Para ambos casos el entrenamiento se realizó de forma supervisada, manteniendo fijos el umbral de confianza y el coeficiente de IoU, el cual describe el grado de superposición de dos boxes, de esta forma cuanto mayor sea la región de superposición, mayor será el coeficiente.

	Red Única	Red Combinada	
		Red HOS	Red Coomassie
<b>mAP</b>	54.7	81.6	88.5
<b>Conf / IoU</b>	0.25 / 0.45	0.25 / 0.45	0.25 / 0.45
<b>N° Params (en millones)</b>	173.1	87.7	87.7

Tabla 1. Comparación de desempeño de los dos esquemas definidos.

- La comparación de desempeño de ambos esquemas se muestra en la Tabla 1. Nótese que al dividir la clasificación en dos partes se obtienen mejores resultados que al emplear una única red con una cantidad similar de parámetros.

- Para mejorar el desempeño de la Red Combinada, se realizaron pruebas con distintos valores para el umbral de confianza y el coeficiente de IoU. Las combinaciones indicadas en rojo en la Figura 2 corresponden a las pruebas para las que se lograron los valores de precisión más elevados.

Figura 2. Resultados de las pruebas de optimización para la red HOS (zq) y para la red Coomassie (dec).

## Conclusiones

- Se desarrolló un modelo basado en aprendizaje profundo para segmentar y clasificar automáticamente imágenes de espermatozoides luego de aplicar las técnicas HOS/Coomassie.

- El modelo construido con las redes entrenadas de forma independiente presentó mejor desempeño que el modelo entrenado con ambas redes de manera conjunta.

- Se desarrolló una herramienta de software con interfaz gráfica que implementa el modelo propuesto, de manera de facilitar su uso.