



Apellidos, nome /Apellidos, nombre: Suárez García, Diego	DNI: [Redacted]	e-mail: [Redacted]	Teléfono de contacto: [Redacted]
Título: Herramienta para evaluar el daño producido en el ADN espermático de centollo a través del análisis de células cometa			Mención cursada: Computación

Resumo / Resumen:

El propósito de mi TFG surge a raíz de un problema real al que se enfrentaba un grupo de biólogos de la Facultad de Ciencias de la UDC. Resumidamente, estos investigadores estaban llevando a cabo un estudio sobre la centolla gallega y necesitaban procesar una serie de imágenes de *células cometa* que obtuvieron de células espermáticas de centollo.

Para poneros un poco en contexto, se denominan células cometa a aquellas que han sido tratadas por el Ensayo Cometa, una técnica de laboratorio que se utiliza para estudiar el nivel de deterioro que se ha haya podido llegar a producir en el ADN de estas. Tras la finalización del ensayo, las células que están sanas mantienen su núcleo intacto, mientras que las que no lo están sufren pérdida de su material nucleico, de forma directamente proporcional a lo dañadas que estén. El proceso concluye con la obtención de imágenes de las células y el procesamiento de estas mediante el uso de herramientas *software* para extraer métricas cuantitativas.

Las herramientas gratuitas a las que el grupo tenía acceso no segmentaban correctamente sus imágenes y tampoco permitían la segmentación manual de estas. En contrapartida, los programas con licencia comercial disponibles eran de coste inasumible.

Así pues, mi trabajo consistió en el análisis y desarrollo de una aplicación de escritorio que permitiese, tanto de forma manual como automática, la segmentación de imágenes de células cometa. En una primera fase y tras analizar el problema a resolver, se estudió el estado del arte en el que se encontraba el *software* en relación con el Ensayo Cometa. El siguiente paso consistió en el análisis de una muestra de las imágenes del equipo y a partir de este desarrollar el algoritmo de segmentación automática mediante el empleo de técnicas de Visión Artificial. En la última fase se construyó la interfaz gráfica y la herramienta de segmentación manual, y se integraron junto al algoritmo de segmentación automática, dando como resultado la aplicación actual.

A grandes rasgos, la aplicación permite al usuario cargar sus imágenes (o bien importándolas de forma independiente o abriendo un proyecto que las contenga previamente guardado), procesarlas y obtener las métricas de cada célula segmentada, viéndolas en la propia aplicación o exportándolas a una hoja de cálculo. También le permite guardar sus imágenes en proyectos distintos para agruparlas y recuperar los resultados de cada sesión sin tener que realizar de nuevo la segmentación.

Las principales tecnologías empleadas en la aplicación fueron *Python* (lenguaje de programación), *OpenCV* (librería de Visión Artificial), *NumPy* (librería de Python para trabajar con *arrays* multidimensionales) y *GTK* (librería de interfaz gráfica), entre otras múltiples.

Posibles aplicaciones / Posibles aplicaciones:

Mi aplicación fue utilizada por el grupo de la Facultad de Ciencias, pudiendo procesar sus imágenes de forma exitosa sin prácticamente tener que utilizar las herramientas de segmentación manual, y sólo desechando las segmentaciones de células erróneas. La previsión es que otras personas descarguen la aplicación una vez publiquen el artículo en el que están trabajando, haciendo referencia a ella.

No obstante, la aplicación no se ciñe exclusivamente a células espermáticas de centollo. Por ejemplo, en los seres humanos se emplea para estudiar cómo nuestros hábitos dietéticos, el tabaquismo o la exposición a determinados metales y/o gases (entre muchos otros), pueden afectar negativamente a nuestro organismo. Esto se puede extrapolar a cualquier otro dominio donde se empleen células eucariotas de cualquier población. Es decir, es aplicable a cualquier área en el que sea necesario estudiar la correlación que pueda existir entre la exposición de organismos a determinadas sustancias tóxicas y el deterioro producido en el ADN de sus células.

Etapas para o seu desenvolvemento futuro / Etapas para su desarrollo futuro:

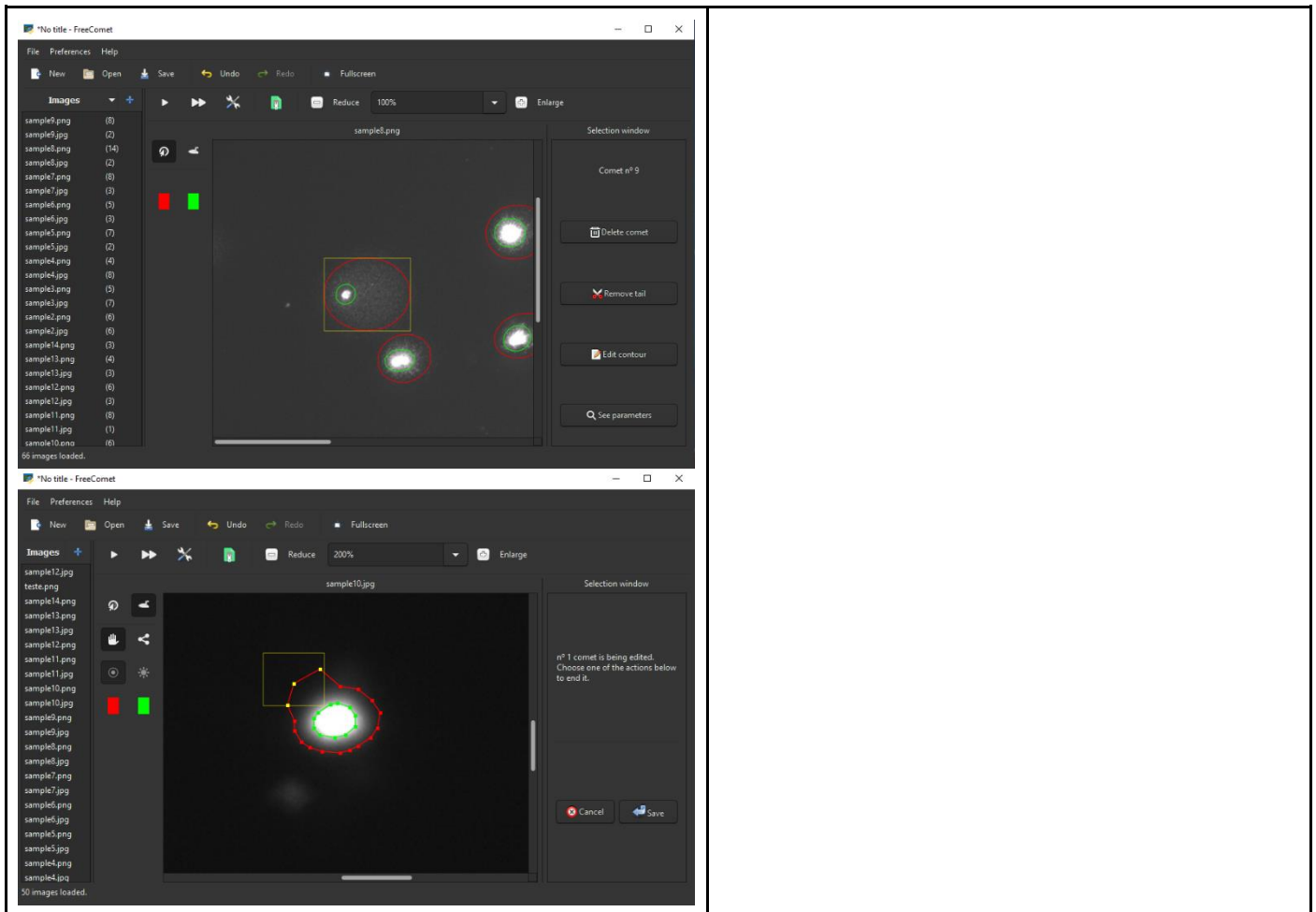
A continuación, listo los puntos más relevantes sobre los que habría que trabajar en una etapa de desarrollo futura:

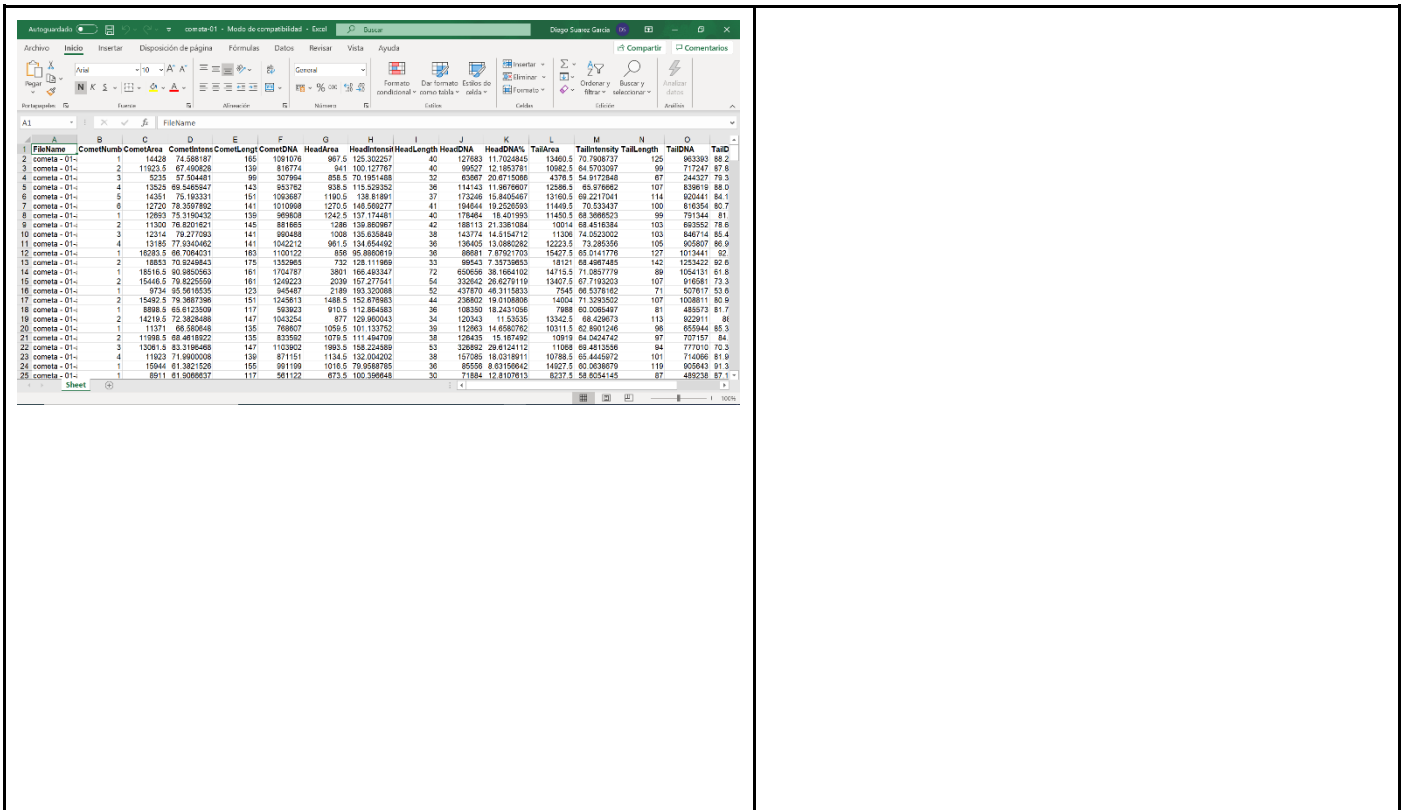
- **Segmentación automática:** el algoritmo de segmentación automática fue diseñado a partir de la muestra del equipo de investigadores de la UDC y, a pesar de las múltiples y distintas formas que llegan a presentar las células en las imágenes, funciona bastante bien. Además, si una célula no es segmentada correctamente, el usuario siempre puede desecharla y/o segmentarla él mismo manualmente. No obstante, las imágenes de células cometa pueden llegar a ser muy diferentes si no se toman de la misma forma (color, relación entre el tamaño de las células con el tamaño global de la imagen, las características de las propias células, ruido en la imagen, ángulo, etc.), por lo que las imágenes que se hayan obtenido en una organización pueden llegar a ser significativamente distintas a las de otra. En un escenario idílico, independientemente del organismo al que pertenezcan los usuarios y de la forma en la que tomen las imágenes, deberían poder segmentarse automáticamente de forma exitosa. Este escenario no es realista, pero conviene intentar que el algoritmo sea lo más genérico posible o al menos incorporar nuevos algoritmos que trabajen mejor con imágenes de otras características. La aplicación ya incluía un segundo algoritmo de segmentación implementado a partir del que se utiliza en *OpenComet*; no obstante, se terminó descartando porque la implementación era muy lenta, pero se empleó para comparar los resultados de ambos algoritmos sobre las imágenes
- **Segmentación manual:** sería conveniente revisar las herramientas y la experiencia que la aplicación ofrece actualmente a los usuarios para segmentar manualmente sus imágenes, para que el proceso sea lo más simple e intuitivo posible.
- **Diseño:** sería conveniente mejorar la experiencia del usuario con la aplicación, mejorando el diseño de las ventanas y redistribuyendo los elementos de la interfaz gráfica para que visualmente la aplicación sea más agradable. Sería muy interesante, con la reciente llegada de la primera versión estable de GTK4, aprovechar la oportunidad para actualizarse y aportarle ese nuevo diseño a la aplicación, estudiando la posibilidad de crear *Widgets* propios para mejorar el desarrollo y mantenimiento de la aplicación.
- **Feedback:** hay que aportarle mayor *feedback* al usuario, tanto de los resultados de las acciones que va ejecutando, como de los posibles errores que ocurran.
- **Logger:** no imprescindible, pero seguiría siendo conveniente para el desarrollador, crear un sistema de *logs* para poder mantener la aplicación.
- **Internacionalización:** actualmente, los textos están traducidos al castellano y al inglés. Habría que adaptarla en

base a los usuarios que empleen la aplicación.

- **Output:** el sistema exporta las métricas en una hoja de cálculo. Sería interesante incluir junto a esta las imágenes con las células segmentadas y numeradas, para que los usuarios puedan referenciar visualmente a qué célula corresponde una métrica sin tener que abrir la aplicación.
- **Tamaño de la aplicación:** la aplicación ocupa entre 200 y 250 MBs, un tamaño excesivo que hay que revisar, posiblemente porque está importando más de lo que debería de las librerías empleadas.
- **Código:** No imprescindible pero recomendable en este punto para su posible futura ampliación y futuro mantenimiento (ya que la aplicación es relativamente pequeña), estaría bien revisar el diseño de las clases y módulos y cómo interactúan entre sí, y realizar los cambios convenientes para que el código sea lo más legible y mantenible posible.

Imaxes representativas / Imágenes representativas:





File Name	ConteNum	ConteArea	ConteLong	ConteDNA	HeadArea	HeadIntens	HeadLength	HeadDNA	HeadDNA%	TailArea	TailIntens	TailLength	TailDNA	TailD	
cometa-01-1	1	14428	74.588197	165	1001076	987.5	125.302297	40	127083	11.7020816	13480.5	70.7689737	125	963393	98.2
cometa-01-2	2	11923.5	67.490829	139	818774	841	100.127787	40	89527	12.1853781	10982.5	64.5705267	99	717247	87.8
cometa-01-3	3	5235	57.304481	99	307994	838.5	70.1951488	32	83087	20.8713986	4378.5	54.9172648	87	244327	76.3
cometa-01-4	4	13629	66.3485847	143	953762	938.5	115.263382	36	114143	11.9876607	12386.5	65.676862	107	838916	86.0
cometa-01-5	5	14251	75.163321	151	1063887	1190.5	138.81891	37	173268	16.8405687	13760.5	69.2217041	114	800441	84.1
cometa-01-6	6	12730	78.397802	141	1010968	1270.5	146.989277	41	184644	19.2526953	11460.5	70.5234377	100	816564	92.7
cometa-01-7	7	17893	73.3150432	139	969808	1242.5	137.174481	40	178484	18.401893	11450.5	68.3698523	99	791344	81
cometa-01-8	8	11530	78.801421	145	891665	1288	139.865987	42	188113	21.3591884	10014	68.4516384	103	692552	78.6
cometa-01-9	9	12314	79.277093	141	860488	1008	135.632848	38	143774	14.5154712	11308	74.0523002	103	846714	85.4
cometa-01-10	10	13185	77.0340462	141	1042212	901.5	134.854492	36	136405	13.0882692	12223.5	73.285359	105	805807	86.9
cometa-01-11	11	16383.5	68.7046331	163	1100120	806	95.808919	36	89891	8.87617703	15477.5	65.5141776	127	1013441	92
cometa-01-12	12	18853	70.8249843	175	1322965	732	128.111869	33	89543	7.35739633	18121	68.4987485	142	1253422	82.6
cometa-01-13	13	18316.5	69.8050563	161	1762187	3801	166.492347	72	606656	28.1684162	14715.5	71.2807779	98	1054131	61.8
cometa-01-14	14	15446.5	78.8225559	161	1249223	2039	157.277541	54	332842	28.8279119	13407.5	67.1193203	107	816581	73.3
cometa-01-15	15	8734	58.5818535	123	345487	2189	193.320388	52	437870	48.3118833	7545	66.5378162	71	507817	53.6
cometa-01-16	16	15467.5	78.3877298	151	1263613	1481.5	150.879883	44	238802	19.0188806	14264	71.5295202	107	1008911	80.9
cometa-01-17	17	8896.5	68.6123509	117	593923	910.5	112.864383	36	108350	18.2431036	7888	60.2085487	81	485873	81.7
cometa-01-18	18	14216.5	72.3826498	147	1043264	877	129.860343	44	120343	11.53035	13342.5	68.426973	113	822911	86
cometa-01-19	19	11371	66.850648	135	789827	1059.5	101.133752	39	112983	14.6588782	10311.5	62.8801246	98	655044	85.3
cometa-01-20	20	1198.5	68.4818522	135	833582	1078.5	111.484709	38	128435	15.187482	10919	68.0424742	97	707157	84
cometa-01-21	21	13061.5	83.198468	147	1103902	1982.5	158.224880	53	328892	29.6124112	11068	69.4813556	94	777010	70.3
cometa-01-22	22	11623	71.8002008	139	871151	1134.5	132.054202	38	167089	18.0218911	10788.5	66.4445972	101	716086	81.9
cometa-01-23	23	15044	81.3811528	155	991199	1016.5	79.9589785	36	86558	8.03159642	14927.5	60.2638679	119	805943	81.3
cometa-01-24	24	8911	61.3056637	117	561122	873.5	100.359648	30	71884	12.8107913	8337.5	58.8034145	87	495238	87.3

V Autorizo a consulta por parte dos membros da comisión evaluadora da memoria do meu proxecto / Autorizo la consulta por parte de los miembros del tribunal de la memoria de mi proyecto.

Instruccions para o depósito da memoria / Instrucciones para el depósito de la memoria:

Débase depositar no OneDrive da UDC, dentro da carpeta co seu nome de usuario incluída en:

6 edición - Premio TFG Aplicado (https://udcgalmys.sharepoint.com/:f/g/personal/nieves_pedreira_udc_es/EqGpB2K77ORHtM3ik-lIGBYBAFqaBS8pZELY9pQHj32DtQ?e=Hl2bC1)

Se debe depositar en el OneDrive de la UDC, dentro de la carpeta con su nombre de usuario incluída en:

6 edición - Premio TFG Aplicado (https://udcgalmys.sharepoint.com/:f/g/personal/nieves_pedreira_udc_es/EqGpB2K77ORHtM3ik-lIGBYBAFqaBS8pZELY9pQHj32DtQ?e=Hl2bC1)