

UNIVERSIDAD AUTÓNOMA DE MADRID

ESCUELA POLITÉCNICA SUPERIOR



Máster en ...

TRABAJO FIN DE MÁSTER

TÍTULO DEL TFM

Autor: Nombre Apellido1 Apellido2

Tutor: Nombre Apellido1 Apellido2

Ponente: Nombre Apellido1 Apellido2

MES 20xx

TÍTULO DEL TFM

Autor: Nombre Apellido1 Apellido2
Tutor: Nombre Apellido1 Apellido2
Ponente: Nombre Apellido1 Apellido2

Grupo de la EPS (opcional)
Dpto. de XXXXX
Escuela Politécnica Superior
Universidad Autónoma de Madrid
MES 20xx

Resumen

Resumen

Palabras Clave

Abstract

Key words

Agradecimientos

Índice general

Índice de Figuras	x
Índice de Tablas	xii
1. Introducción	1
1.1. Motivación del proyecto	1
1.2. Objetivos y enfoque	2
1.3. Metodología y plan de trabajo	2
1.3.1. Metodología	2
1.3.2. Plan de Trabajo	2
2. Reconocimiento de iris. Estado del arte	5
2.1. Introducción	5
2.2. Historia, nacimiento y evolución.	5
2.3. La anatomía del ojo	5
2.3.1. Aspectos diferenciadores del iris	5
2.4. Adquisición del Iris	5
2.4.1. Introducción	5
2.4.2. Esquemas de adquisición tradicionales	5
2.4.3. Consideraciones sobre la iluminación	5
2.4.4. Posicionamiento del Iris	5
2.4.5. Sistemas comerciales de adquisición	5
2.5. Localización y segmentación del Iris	5
2.5.1. Introducción	6
2.5.2. Metodología de J. Daugman y derivadas	6
2.5.3. Metodología de R. Wildes y derivadas	6
2.5.4. Otras metodologías	6
2.5.5. Comparativa de metodologías	6
2.5.6. Detección de pestañas y ruido	6
2.6. Normalización del tamaño	6
2.6.1. Daugman's Rubber Sheet Model	6

2.6.2.	Image Registration	6
2.6.3.	Normalización en ángulo	6
2.6.4.	Mejora del contraste y eliminación de ruido	6
2.7.	Algoritmos de Codificación	6
2.7.1.	Metodología de Daugman: Filtros de Gabor	6
2.7.2.	Metodologías alternativas a la de Daugman	6
2.7.3.	Metodologías de Wildes. Vectores de características reales (no binarios)	6
2.8.	Algoritmos de Matching	6
2.8.1.	Introducción	6
2.8.2.	Distancia de Hamming	6
2.8.3.	Distancia euclídea ponderada	6
2.8.4.	Correlación normalizada	6
2.9.	Problemática y retos futuros	6
2.9.1.	Segmentación	7
2.9.2.	Captura ideal no invasiva	7
2.10.	Competiciones o Evaluaciones de Iris	7
2.10.1.	The Iris Challenge Evaluation (ICE)	7
2.10.2.	The Noisy Iris Challenge Evaluation (NICE)	7
2.11.	Bases de datos	7
2.11.1.	CASIA	7
2.11.2.	BioSec Baseline y BioSecurID	7
3.	Sistema, diseño y desarrollo	9
3.1.	Segmentación	9
3.2.	Normalización	9
3.3.	Codificación	9
3.4.	Matching	9
4.	Experimentos Realizados y Resultados	11
4.1.	Bases de datos y protocolo	11
4.2.	Sistemas de referencia	11
4.3.	Escenarios de pruebas	11
4.4.	Experimentos del sistema completo	11
5.	Conclusiones y trabajo futuro	13
	Glosario de acrónimos	15
	Bibliografía	16

A. Manual de utilización	19
B. Manual del programador	21

Índice de Figuras

1.1. Ejemplo pie de figura 1	1
--	---

Índice de Tablas

1

Introducción

1.1. Motivación del proyecto

Ejemplo de referencia a la bibliografía [1].

Ejemplo de imagen:



Figura 1.1: Ejemplo pie de figura 1

1.2. Objetivos y enfoque

1.3. Metodología y plan de trabajo

1.3.1. Metodología

1.3.2. Plan de Trabajo

Sprint 1 (18/9/2018 - 3/10/2018)

El primer sprint ha estado centrado tanto en definir con más exactitud la dirección del proyecto como en un primer acercamiento a las principales herramientas con las que va a desarrollarse.

Tras unos primeros pasos con Galaxy [2] y Docker [3], se ha tomado como referencia el trabajo Bioinfworkflow de Sergio Chico [4] como base para la imagen Docker del proyecto. Dado que el proyecto de Github daba algunos problemas en la instalación, se ha desarrollado un script propio que produce los mismos resultados.

Una vez se ha tenido disponible la imagen de Docker, el sprint se ha centrado en algunos aspectos importantes para partes futuras del desarrollo. Entre ellos destaca la investigación acerca del formato de los workflows de Galaxy (.ga) ya que en un futuro será necesario generar este tipo de ficheros para introducirlos en Galaxy. También resulta relevante la investigación acerca de las posibilidades que ofrece la API de Galaxy [4] y su utilidad en Bioblend [5], que nos facilitan la opción de utilizar Galaxy sin necesidad de hacerlo a través de su interfaz.

Sprint 2 (4/10/2018 - 17/10/2018)

La primera semana de este sprint ha estado dirigida a lograr una imagen Docker de Galaxy que contenga un set de herramientas básicas para formar un primer workflow. Se han valorado varias opciones de instalación en las que se han utilizado tanto la imagen básica de Galaxy [6] como la imagen de Bioinfworkflow [4]. Finalmente se ha optado por utilizar Bioinfworkflow ya que parte de las herramientas necesarias ya estaban incluidas. Para realizar esta tarea se ha creado un nuevo fichero Dockerfile así como el listado de herramientas necesarias para su instalación.

Sprint 3 (18/10/2018 - 31/10/2018)

El sprint ha estado centrado en la correcta ejecución del workflow con las herramientas iniciales desde Galaxy. Durante el proceso de configuración han surgido varias complicaciones que han impedido terminar el workflow completo en este sprint. En un principio han surgido problemas con el filtrado de calidad utilizando Prinseq. Este problema no ha llegado a ser resuelto en este sprint a falta de tratar el tema con el grupo de Tecnofood. A continuación se encontraron ciertos problemas con el formato de salida de la herramienta Prokka. A pesar de que la salida está marcada como formato gff3, un parámetro interno lo etiquetaba como gff. Esto impedía que la salida de Prokka fuese introducida como input en las herramientas siguientes.

Dados estos errores, se decidió trabajar en paralelo con la API de Galaxy desde Python para intentar ejecutar tanto las herramientas como el workflow de una manera menos restringida. Finalmente se ha llevado a cabo el desarrollo necesario para subir los ficheros a un historial y ejecutar cada una de las herramientas del workflow desde Python.

Sprint 4 (1/11/2018 - 14/11/2018)

La prioridad en este punto se ha centrado en completar el workflow desde Galaxy. Han surgido varios problemas en esta tarea. La primera es un bug en Mac por el cual los archivos eliminados dentro de Docker no se eliminan del todo y quedan fijados en un fichero residual. Esto implica que cada cierto tiempo hay que eliminar la imagen completa de Docker para poder liberar espacio, dado el gran tamaño de los archivos con los que se trabaja. Debido a ello, la tarea de completar el workflow se ha visto retrasada. Además, la ejecución de la herramienta Roary a través de Galaxy ejecuta sin errores pero no devuelve ningún resultado, lo que ha impedido continuar con la parte final del workflow.

Además de la tarea ya comentada, en este sprint se ha trabajado en el acceso al contenedor Docker desde otro ordenador en una red local. Con el objetivo de desplegar el servicio en un servidor.

También se ha tomado contacto con la interfaz gráfica, realizando unas pruebas en las que simplemente se muestra alguna información extraída de la API de Galaxy en etiquetas creadas con PyQt.

Sprint 5 (15/11/2018 - 28/11/2018)

Al igual que en los sprints anteriores, gran parte de la carga de trabajo se ha centrado en resolver problemas en la ejecución del workflow a través de Galaxy. Se han realizado numerosas ejecuciones para comprobar si las salidas de cada herramienta eran las correctas. Esto ha servido para concluir que, al parecer, un fallo en la herramienta Roary incluida en Galaxy, impide que los ficheros retornados tengan contenido. Esto se ha comprobado a través de la ejecución de Roary standalone con los mismos datos de entrada y los mismos parámetros, obteniendo de esta manera los ficheros correctos.

Para ahorrar en tiempos de ejecución se han utilizado los ficheros fasta ya generados previamente, no los generados con la herramienta Spades de nuestro propio workflow. En el próximo sprint se tratará de integrar la parte previa a los pasos que ya son correctos.

También se ha añadido un fichero .gitignore para evitar la existencia de ficheros irrelevantes en el repositorio.

Parte del trabajo de este sprint se ha destinado a la redacción de la propuesta de proyecto y de la introducción a la documentación del mismo.

2

Reconocimiento de iris. Estado del arte

2.1. Introducción

2.2. Historia, nacimiento y evolución.

2.3. La anatomía del ojo

2.3.1. Aspectos diferenciadores del iris

2.4. Adquisición del Iris

2.4.1. Introducción

2.4.2. Esquemas de adquisición tradicionales

2.4.3. Consideraciones sobre la iluminación

2.4.4. Posicionamiento del Iris

2.4.5. Sistemas comerciales de adquisición

2.5. Localización y segmentación del Iris

2.5.1. Introducción

2.5.2. Metodología de J. Daugman y derivadas

2.5.3. Metodología de R. Wildes y derivadas

2.5.4. Otras metodologías

2.5.5. Comparativa de metodologías

2.5.6. Detección de pestañas y ruido

2.6. Normalización del tamaño

2.6.1. Daugman's Rubber Sheet Model

2.6.2. Image Registration

2.6.3. Normalización en ángulo

2.6.4. Mejora del contraste y eliminación de ruido

2.7. Algoritmos de Codificación

2.7.1. Metodología de Daugman: Filtros de Gabor

2.7.2. Metodologías alternativas a la de Daugman

Filtros Log-Gabor

Wavelets

Haar Wavelet

Transformada Discreta del Coseno (DCT)

2.7.3. Metodologías de Wildes. Vectores de características reales (no binarios)

2.8. Algoritmos de Matching

2.8.1. Introducción

2.8.2. Distancia de Hamming

2.8.3. Distancia euclídea ponderada

2.8.4. Correlación normalizada

2.9. Problemática y retos futuros

2.9.1. Segmentación

2.9.2. Captura ideal no invasiva

2.10. Competiciones o Evaluaciones de Iris

2.10.1. The Iris Challenge Evaluation (ICE)

2.10.2. The Noisy Iris Challenge Evaluation (NICE)

2.11. Bases de datos

2.11.1. CASIA

2.11.2. BioSec Baseline y BioSecurID

3

Sistema, diseño y desarrollo

3.1. Segmentación

3.2. Normalización

3.3. Codificación

3.4. Matching

4

Experimentos Realizados y Resultados

4.1. Bases de datos y protocolo

4.2. Sistemas de referencia

4.3. Escenarios de pruebas

4.4. Experimentos del sistema completo

5

Conclusiones y trabajo futuro

Glosario de acrónimos

- **IS**: Iris Subject
- **DCT**: Discrete Cosine Transform
- **WED**: Weighted Euclidean Distance

Bibliografía

- [1] Autor Apellidos. Título del artículo. *Revista de publicación*, pages 65–73, 2008.
- [2] Galaxy. <https://usegalaxy.org/>.
- [3] Docker - Build, Ship, and Run Any App, Anywhere. <https://www.docker.com/>.
- [4] Sergio Chico. :whale: Docker with Galaxy for Bioinformatic Bacterial Sequencing Workflows: Serux/docker-galaxy-BioInfWorkflow. <https://github.com/Serux/docker-galaxy-BioInfWorkflow>, June 2018. original-date: 2018-05-17T01:10:05Z.
- [5] Galaxy API. <https://galaxyproject.org/develop/api/>.
- [6] Björn Gruening. :whale::bar_chart::books: Docker Images tracking the stable Galaxy releases.: bgruening/docker-galaxy-stable. <https://github.com/bgruening/docker-galaxy-stable>, October 2018. original-date: 2014-08-12T13:26:14Z.



Manual de utilización



Manual del programador