

Universidad
Rey Juan Carlos

Escuela Técnica Superior
de Ingeniería Informática

Grado en Ingeniería del Software

Curso 2023-2024

Trabajo Fin de Grado

**MÉTODOS DE APRENDIZAJE AUTOMÁTICO
PARA LA CLASIFICACIÓN DE
COMPORTAMIENTOS ESTEREOTIPADOS**

Autor: Gonzalo Ortega Carpintero
Tutor académico: Alberto Fernández Gil
Tutora CSIC: Paula Peixoto Moledo

Agradecimientos

(Por redactar)

Resumen

(Por redactar)

Palabras clave:

- Python
- Neurociencia
- Aprendizaje automático
- ...

Índice de contenidos

Índice de tablas	X
Índice de figuras	XII
Índice de códigos	XIV
Introducción	1
Obejetivos	2
1. Fundamentos teóricos	3
1.1. Métodos no supervisados	3
1.1.1. Agrupación por afinidad	3
1.1.2. Agrupación aglomerada	3
1.1.3. Análisis de componentes principales	3
1.2. Métodos supervisados	4
1.2.1. Red neuronal	4
2. Análisis y experimentación	5
2.1. Preprocesado de datos	5
2.1.1. DeepLabCut	5
2.1.2. Filtrado e interpolación	7
2.1.3. Computo de variables a analizar	8
2.2. Procesado de datos	9
2.2.1. Matriz de similitud	9
2.2.2. Agrupamiento por afinidad	9
2.2.3. Reducción de dimensionalidad para visualización	9
2.2.4. Entrenamiento de una red neuronal	9
3. Resultados	10
4. Conclusiones y trabajos futuros	11
Bibliografía	11

Apéndices	14
A. Este es el primer apéndice	16
A.1. Ejemplo de sección	16

Índice de tablas

2.1. Extracto del DataFrame de Pandas de los datos sin procesar de DeepLabCut.	6
---	---

Índice de figuras

2.1.	Salida de DeepLabCut del animal 4128 el 02-12-2020, 1:00:37. . .	5
2.2.	Triangulaciones de los puntos del animal 4128 el 02-12-2020, 1:00:37.	6
2.3.	Trayectoria de un punto de cabeza del animal 4128 el 02-12-2020.	7
2.4.	Trayectoria filtrada de un punto de la cabeza del animal.	7
2.5.	Trayectoria interpolada de un punto de la cabeza del animal. . . .	8

Índice de códigos

Introducción

(Provisional)

La Neurociencia es la disciplina científica que estudia el sistema nervioso y todas sus componentes. Para el estudio del funcionamiento del cerebro pueden utilizarse multitud de técnicas diferentes, pero para evitar la intrusión en el sujeto de estudio de muchas de ellas, muchos experimentos e investigaciones se basan en el estudio del comportamiento. Es la etología la encargada de este tipo de estudio en animales y en la cual, históricamente, la categorización de comportamientos dependía fuertemente de la observación manual. Eso ha conllevado siempre una excesiva cantidad de tiempo y las desventajas de los posibles errores de percepción humanos, limitando la cantidad y la calidad de los análisis a realizar.

Gracias a los recientes avances en algoritmos de aprendizaje automático, hoy en día es posible computar multitud de datos de forma simultánea, agilizando y automatizando análisis como el ilustrado previamente.

Este Trabajo de Fin de Grado se ha realizado a lo largo de una estancia de prácticas académicas en el Jercog's Team, un equipo de investigación perteneciente al Instituto Cajal, instituto de Neurociencia del Congreso Superior de Investigaciones Científicas. El equipo está centrado en el estudio de la memoria mediante experimentos con ratones, y uno de sus proyectos abiertos consistía en poder elaborar una herramienta para clasificar automáticamente los comportamientos estereotipados de los ratones. Estos son comportamientos cortos, repetitivos y con cierta tendencia a generar patrones, tales como rascarse, caminar en círculos o levantarse a dos patas, ejecutados sin ninguna finalidad, y usualmente inducidos por estar en entornos cerrados y artificiales.

Durante la estancia se han valorado multitud de técnicas de análisis y procesamiento de datos, haciendo hincapié en métodos de aprendizaje automático para tratar de completar el proyecto gracias a los últimos avances en computación.

(Para realizar este trabajo se han consultado los libros [1, Introduction to Machine Learning with Python] [2, Deep Learning with PyTorch] y se han tomado ideas y procedimientos de los artículos [3] y más).

Objetivos

(Provisional)

El objetivo de este trabajo es hacer una breve introducción a ciertas técnicas de aprendizaje automático aplicables al procesado de datos en el análisis de experimentos con animales. Para ello el trabajo se divide en dos capítulos principales.

En el primero abordaremos el estudio teórico de ciertos algoritmos de agrupamiento, reducción de dimensionalidad y de clasificación supervisada. Todo ello desde un punto de vista teórico con ejemplos fáciles de visualizar para ayudar al lector a comprender en profundidad los métodos.

En el segundo capítulo, desarrollaremos paso a paso el análisis real ejercido durante la estancia en el Instituto Cajal. El principal objetivo durante estos meses ha sido buscar la mejor forma de, dados una serie de videos de ratones en su caja, detectar automáticamente diferentes tipos de comportamientos. El fin último del proyecto es estudiar la posibilidad de clasificar dos grupos de animales, control y medicados, basándose en los comportamientos estereotipados.

Para el preprocesado de los videos se ha utilizado la herramienta DeepLabCut, que mediante el uso de redes neuronales y una relativamente pequeña muestra de entrenamiento computa una estimación de la posición de los animales en cada uno de los fotogramas. Tras el preprocesado haremos una revisión de los artículos científicos que han motivado las diferentes técnicas de agrupamiento que hemos acabado utilizando, y finalizaremos desarrollando cada una de ellas a la par que desglosando el código en Python utilizado en cada apartado.

1

Fundamentos teóricos

1.1. Métodos no supervisados

1.1.1. Agrupación por afinidad

1.1.2. Agrupación aglomerada

1.1.3. Análisis de componentes principales

1.2. Métodos supervisados

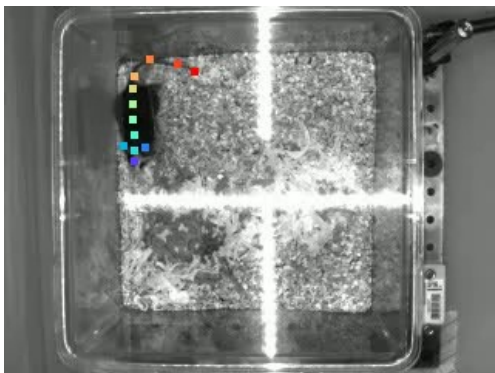
1.2.1. Red neuronal

2

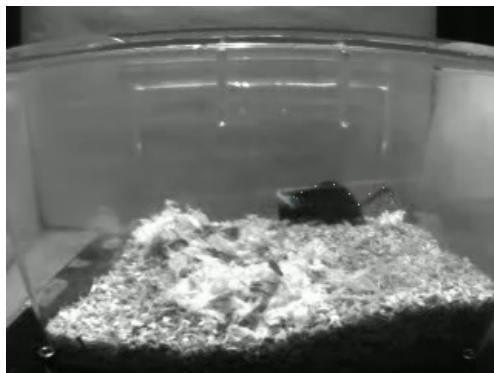
Análisis y experimentación

2.1. Preprocesado de datos

2.1.1. DeepLabCut



(a) Salida de video cenital.

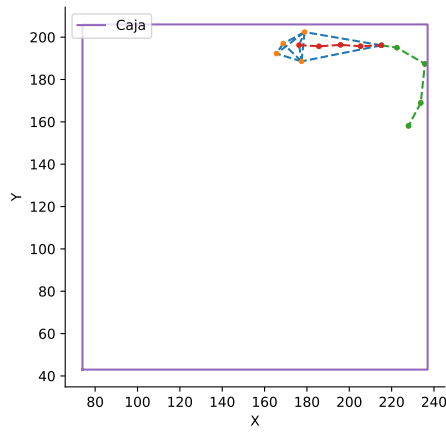


(b) Salida de video lateral.

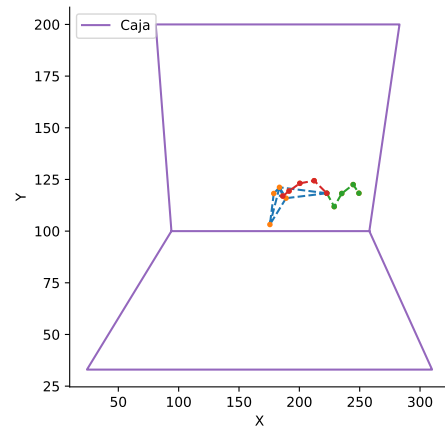
Figura 2.1: Salida de DeepLabCut del animal 4128 el 02-12-2020, 1:00:37.

	Nosex	Nosey	Noselikelihood	Headx	Heady	...
0	136.165344	177.722496	0.000084	129.790253	174.772552	...
1	162.032005	201.444756	0.942181	168.152061	202.420639	...
2	156.297043	200.326378	0.000073	162.436028	203.156837	...
3	155.370415	199.043297	0.000277	159.507599	200.199928	...
4	149.272644	197.677170	0.000045	155.493912	198.814835	...
...

Tabla 2.1: Extracto del DataFrame de Pandas de los datos sin procesar de DeepLabCut.



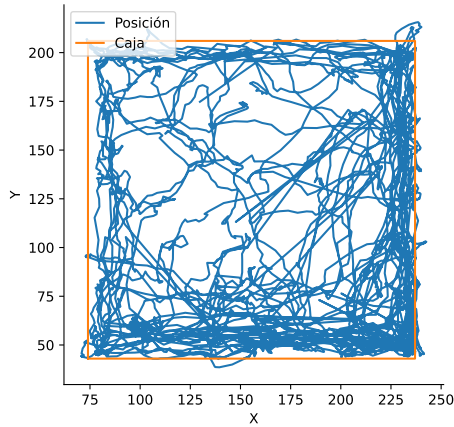
(a) Vista cenital.



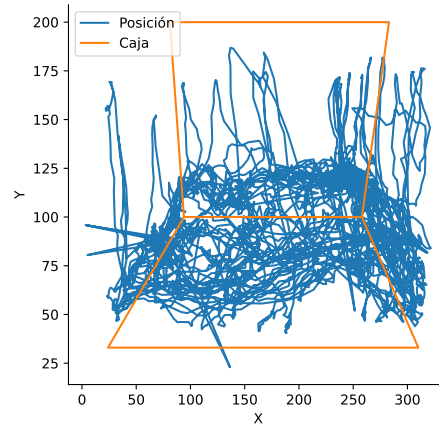
(b) Vista lateral.

Figura 2.2: Triangulaciones de los puntos del animal 4128 el 02-12-2020, 1:00:37.

2.1.2. Filtrado e interpolación

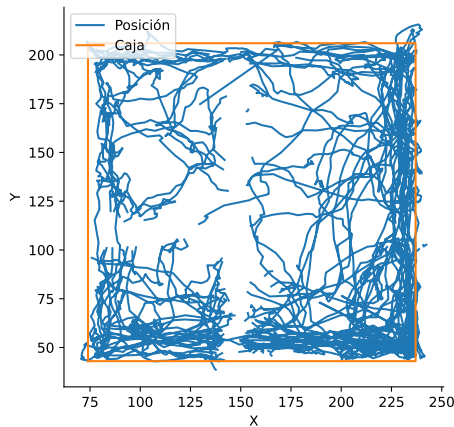


(a) Vista cenital.

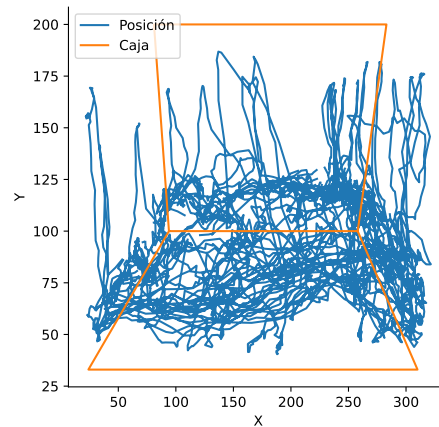


(b) Vista lateral.

Figura 2.3: Trayectoria de un punto de cabeza del animal 4128 el 02-12-2020.

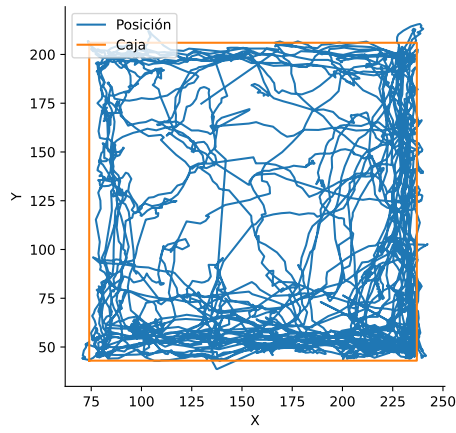


(a) Vista cenital.

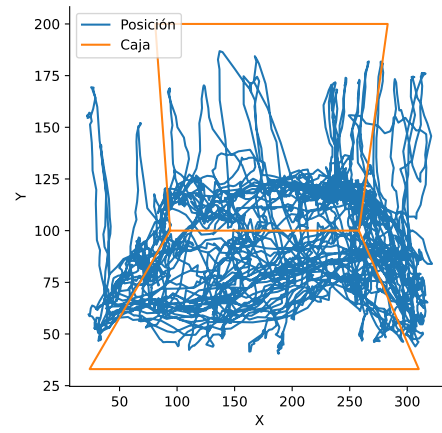


(b) Vista lateral.

Figura 2.4: Trayectoria filtrada de un punto de la cabeza del animal.



(a) Vista cenital.



(b) Vista lateral.

Figura 2.5: Trayectoria interpolada de un punto de la cabeza del animal.

2.1.3. Computo de variables a analizar

2.2. Procesado de datos

2.2.1. Matriz de similitud

2.2.2. Agrupamiento por afinidad

2.2.3. Reducción de dimensionalidad para visualización

2.2.4. Entrenamiento de una red neuronal

3

Resultados

(Por redactar)

4

Conclusiones y trabajos futuros

(Por redactar)

Bibliografía

- [1] A. C. Müller and S. Guido, *Introduction to Machine Learning with Python*. O'Reilly, 2017.
- [2] E. Stevens, L. Antiga, and T. Viehmann, *Deep Learning with PyTorch*. Manning, 2020.
- [3] A. Mathis, P. Mamidanna, K. M. Cury, T. Abe, V. N. Murthy, M. W. Mathis, and M. Bethge, “Deeplabcut: markerless pose estimation of user-defined body parts with deep learning,” *Nature Neuroscience*, 2018. [Online]. Available: <https://doi.org/10.1038/s41593-018-0209-y>

Apéndice



Este es el primer apéndice

A.1. Ejemplo de sección

Sección del apéndice