

ESCUELA TÉCNICA SUPERIOR DE INGENIEROS
INDUSTRIALES Y DE TELECOMUNICACIÓN

UNIVERSIDAD DE CANTABRIA



Trabajo Fin de Grado

**Identificación de especies a través del
análisis de la composición ósea mediante
espectroscopía LIBS**

**(Species identification through bone
composition analysis using LIBS
spectroscopy)**

Para acceder al Título de

***Graduado en
Ingeniería de Tecnologías de Telecomunicación***

Autor: Mario Sánchez García

Septiembre – 2025

GRADUADO EN INGENIERÍA DE TECNOLOGÍAS DE TELECOMUNICACIÓN

CALIFICACIÓN DEL TRABAJO FIN DE GRADO

Realizado por: Mario Sánchez García

Director del TFG: Jesús María Mirapeix Serrano

Título: “Identificación de especies a través del análisis de la composición ósea mediante espectroscopía LIBS ”

Title: “Species identification through bone composition análisis using LIBS spectroscopy “

Presentado a examen el día: 18 de septiembre de 2025

para acceder al Título de

GRADUADO EN INGENIERÍA DE TECNOLOGÍAS DE TELECOMUNICACIÓN

Composición del Tribunal:

Presidente (Apellidos, Nombre): Mañana Canteli, Mario

Secretario (Apellidos, Nombre): Irastorza Teja, José Ángel

Vocal (Apellidos, Nombre): Cobo García, Adolfo

Este Tribunal ha resuelto otorgar la calificación de:

Fdo: El Presidente

Fdo: El Secretario

Fdo: El Vocal

Fdo: El Director del TFG

(sólo si es distinto del Secretario)

Vº Bº del Subdirector

Trabajo Fin de Grado Nº

(a asignar por Secretaría)

Agradecimientos

Quiero agradecer a mi familia; a mis padres por apoyarme siempre, confiar en mí y educarme en la humildad y la constancia para conseguir mis objetivos. A mi hermana por hacer el camino más ameno. A mis abuelos por estar presentes y pendientes en cada paso que doy. Y a mi pareja, por su paciencia, cariño y apoyo, que han hecho que esta etapa fuera más llevadera.

A mis compañeros de carrera, en especial a los que desde el primer día se han convertido en grandes amigos para toda la vida.

Y a los profesores, por haber estado dispuestos a enseñarme y ayudarme a mejorar en todo momento. En especial a Jesús Mirapeix por estar pendiente de mí y ayudarme durante todo el año con este trabajo.

Acrónimos

LIBS: Laser Induced Breakdown Spectroscopy.

Nd:YAG: Neodymium-doped Yttrium Aluminum Garnet.

ICP-OES: Inductively Coupled Plasma Optical Emission Spectroscopy.

XRF: X-ray Fluorescence.

ADN: Ácido desoxirribonucleico

CF-LIBS: Calibration-Free Laser-Induced Breakdown Spectroscopy

IMO: Índice Medular Óseo

LCM: Longitud del Canal Medular

LH: Longitud del Hueso

ZooMS: Zooarcheology by Mass Spectrometry

MALDI-TOF MA: Matrix-Assisted Laser Desorption/Ionization – Time of Flight Mass Spectrometry

PCR: Polymerase Chain Reaction

SVM: Support Vector Machine

kNN: k-Nearest Neighbors

Lista de figuras

Figura 1. Descomposición de luz solar en diferentes colores tras pasar por un prisma....	13
Figura 2. Longitud de onda	14
Figura 3. Espectro de la luz solar	14
Figura 4. Espectro de absorción de algunos gases	15
Figura 5. Espectro de emisión de algunos elementos	15
Figura 6. Emisión de luz	16
Figura 7. Absorción de luz	16
Figura 8. Espectroscopía de ruptura inducida por láser	18
Figura 9. Espectro de emisión para una muestra de hueso	19
Figura 10. Sistema LIBS portátil para el análisis de pesticidas en plantas	20
Figura 11. Comparación entre ICP-OES y LIBS para Mg	21
Figura 12. Monedas romanas analizadas	22
Figura 13. Correlación entre las medidas mediante LIBS y XRF	22
Figura 14. Concentración de Mg/Ca analizada con LIBS	23
Figura 15. Espectro recogido tras el análisis de una muestra de suelo	24
Figura 16. Análisis mediante LIBS de las muestras óseas	25
Figura 17. Análisis de los elementos usando el modelo que emplea a) grupo 1, b) grupo 2, c) ambos grupos	26
Figura 18. Espectro normalizado obtenido con LIBS para una muestra ejemplo	27
Figura 19. Espectro normalizado obtenido con LIBS para una muestra ejemplo	27
Figura 20. Espectro LIBS típico de una muestra de hueso	28
Figura 21. Medidas tomadas en hueso bovino	29
Figura 22. Relación entre nivel del canal medular y edad. Regresión lineal	30
Figura 23. Especies analizadas en el estudio	31
Figura 24. Espectros para muestras óseas de conejo y camello	31
Figura 25. Muestras de huesos utilizadas para el estudio	32
Figura 26. Identificación de especies presentes en cada reacción	33
Figura 27. Esquema del sistema LIBS utilizado	34
Figura 28. Láser utilizado para la experimentación	35
Figura 29. Lentes para la focalización del láser hacia la muestra	35
Figura 30. Soporte motorizado	36
Figura 31. Lentes para la recolección de la luz generada	36
Figura 32. Cámara digital para la visualización de la muestra	37
Figura 33. Control remoto	37
Figura 34. Aprendizaje supervisado.....	38

Figura 35. Aprendizaje no supervisado	39
Figura 36. Ejemplo de árbol de decisión	40
Figura 37. Hiperplano separador y vectores de soporte.....	42
Figura 38. Esquema de red neuronal	44
Figura 39. Ejemplo de media por mapeo	45
Figura 40. Ejemplo de media lineal	46
Figura 41 . Huesos de cordero y cabra enterrados	46
Figura 42 . Huesos de cordero y cabra tras ser recogidos y preparados para su análisis en el laboratorio.....	47
Figura 43 . Hueso de ciervo rojo recogido en el monte	47
Figura 44 . Muestras CR1 y CR2 extraídas del cráneo de ciervo rojo	48
Figura 45 . Muestras desconocidas (D) recogidas en el monte	48
Figura 46. Muestras desconocidas (Z) recogidas en el monte.....	48
Figura 47 . Muestras desconocidas de vertebra (V) recogidas en el monte	49
Figura 48. Efecto provocado por LIBS a la muestra C1	50
Figura 49. Evolución del espectro en un mismo punto con 15 disparos para el hueso de cordero 1	51
Figura 50. Espectro obtenido con LIBS en el hueso de cordero 1	53
Figura 51. Espectro obtenido con LIBS en el hueso de cordero 2	53
Figura 52. Espectro obtenido con LIBS en el hueso de cabra	54
Figura 53. Espectro obtenido con LIBS en el hueso de ciervo 1	54
Figura 54. Espectro obtenido con LIBS en el hueso de ciervo 2.....	55
Figura 55. Pestaña de apps de Matlab	56
Figura 56. Pestaña de Classification Learner	57
Figura 57. Pestaña de selección de datos para el entrenamiento	58
Figura 58. Modelos seleccionables para el entrenamiento.....	59
Figura 59. Resultados del entrenamiento	60
Figura 60. Matriz de confusión en el entrenamiento del árbol de decisión	61
Figura 61. Parámetros establecidos para la red neuronal	62
Figura 62. Precisiones de los modelos SVM y árbol de decisión distinguiendo entre individuos	62
Figura 63. Matriz de confusión durante el entrenamiento con el modelo de árbol de decisión	63
Figura 64. Matriz de confusión durante el entrenamiento con el modelo de SVM	63
Figura 65. Espectro obtenido para el hueso desconocido V1.....	64
Figura 66. Espectro obtenido para el hueso desconocido Z1	65

Figura 67. Espectro obtenido para el hueso desconocido O1	66
---	----

Resumen

La finalidad de este Trabajo de Fin de Grado es evaluar el rendimiento de diferentes modelos, basados en *machine learning*, para la clasificación de muestras óseas en las diferentes especies, en función de los espectros obtenidos de ellas, utilizando para ello la técnica de espectroscopía de plasma inducido por láser (LIBS).

Los espectros mencionados, son específicos para cada especie, ya que dependen de los elementos contenidos en cada muestra. Para ello se ha utilizado la herramienta Matlab tanto para guardar y tratar los datos, como para la evaluación y selección del modelo.

A través de los datos obtenidos de muestras óseas de cabra, cordero y ciervo (venado) se ha comparado el rendimiento de los diferentes modelos evaluados a la hora de realizar una clasificación correcta de las muestras en base a sus espectros.

Palabras clave: Algoritmo, Espectro, Espectroscopía de Plasma Inducido por Láser (LIBS), Redes Neuronales, Muestras Óseas, Aprendizaje supervisado.

Abstract

The purpose of this project is to assess the efficiency of different prototypes, based on a machine learning algorithm capable of classifying bone sample spectra from different species based on their elemental composition, thanks to Laser-Induced Breakdown Spectroscopy (LIBS), a technique that captures the unique spectral fingerprint of each sample.

Since the elemental makeup of a bone sample is species-specific, the resulting spectra serve as reliable identifiers. MATLAB was used as the primary analytical tool—not only for data storage and preprocessing, but also for building the classification prototype

Through the data obtained from goat, lamb, and deer bone samples, the performance of the different models was compared in terms of their ability to correctly classify the samples based on their spectra.

Keywords: Algorithm, Spectra, Laser-Induced Breakdown Spectroscopy (LIBS), Neural Networks, Bones Samples, Supervised Learning.

Índice

Acrónimos	5
Lista de figuras.....	6
Resumen	9
Abstract.....	9
Objetivo y contexto.....	12
Objetivo	12
Contexto.....	12
1. Introducción a la espectroscopía.....	13
1.1. Desarrollo histórico	13
1.2. Tipos de espectroscopía	14
2. Espectroscopía de ruptura inducida por láser (LIBS)	18
2.1. Funcionamiento de los sistemas LIBS	18
2.2. Ejemplos de estudios realizados con LIBS	20
2.3. Comparación de LIBS con técnicas espectroscópicas alternativas	24
3. Estudios centrados en el análisis de huesos	25
4. Montaje del sistema LIBS utilizado.....	34
5. Modelos de clasificación y redes neuronales.....	38
5.1. Introducción al <i>machine learning</i>	38
5.2. Tipos de modelos de clasificación en Matlab.....	39
5.3. Elección del modelo utilizado	44
6. Desarrollo de un programa en Matlab para la distinción de especies usando los espectros obtenidos con LIBS.	45
6.1. Introducción y objetivo del desarrollo.....	45
6.2. Medidas realizadas.....	49
6.3. Desarrollo del programa	52
6.3.1. Almacenamiento de las medidas y representación de los espectros	52
6.3.2. Separación de los datos en entrenamiento y validación y creación de las matrices asociadas.....	55
6.3.3. Diseño e implementación de la red neuronal.....	56
6.3.4. Diferenciación entre individuos de la misma especie.....	62
6.3.5. Clasificación de los huesos desconocidos	64
6.3.6. Conclusión y selección de modelo	66
7. Conclusión y líneas futuras	67
Conclusión	67
Líneas futuras	67

Anexo 1: Código Matlab utilizado	69
Referencias	76

Objetivo y contexto

Objetivo

En este trabajo de fin de grado, se van a analizar huesos de diferentes especies animales utilizando la técnica de espectroscopía de ruptura inducida por láser (*Laser Induced Breakdown Spectroscopy*: LIBS). Para ello, se disparará con el láser LIBS a las muestras, obteniendo así el espectro de emisión del fragmento y tratando de conocer su composición elemental. De esta forma, en función de las diferencias entre los espectros obtenidos para cada especie y, por lo tanto, de los elementos que lo compongan, podremos tratar de establecer una metodología que permita clasificar cada muestra con su correspondiente especie en base a sus espectros de emisión obtenidos al utilizar la técnica LIBS.

Contexto

La técnica de espectroscopía de ruptura inducida por láser es una técnica muy común en el análisis de muestras, ya que aporta mucha información de la sección analizada, sin apenas dañarla y con láseres no demasiado voluminosos ni costosos, por lo que es una técnica óptima para realizar estudios en prácticamente cualquier tipo de muestra.

Como punto de partida para este trabajo de fin de grado, se han tomado diversos artículos de investigación, en los que también se utiliza la técnica LIBS para realizar distintos estudios en huesos, asentando con ello la posibilidad de realizar el objetivo buscado en este TFG. Principalmente se han tenido en cuenta los 3 siguientes artículos al ser los que trabajan en aplicaciones de análisis de huesos usando LIBS.

En primer lugar, se ha tomado el trabajo de investigación de, Marín Roldán et Al (2020) en el cual utilizan LIBS para el análisis de huesos de fauna, consiguiendo discriminar y ensamblar huesos de diferentes especímenes de ciervos de una misma especie [1].

Por otro lado, en 2024, Doaa Youssef, Vincent Motto-Ros y Mohamed Abdel-Harith, tratan de realizar un programa basado en *machine learning* que discrimine huesos humanos arqueológicos, logrando desarrollar un sistema que identifica restos de huesos humanos mezclados con precisión y minimizando el tiempo y almacenamiento utilizado [2].

Por último, para terminar de asentar las bases del trabajo de fin de grado, se ha utilizado la investigación realizada en 2013 por Kasem, M. A., Gonzalez, J. J., Russo, R. E., Harith, y M. A, que utilizaron LIBS para analizar huesos de animales y estudiar la viabilidad de la técnica para estudios forenses. Llegando a la conclusión de que haciendo una calibración simple con una muestra de referencia, era posible obtener resultados interesantes. Además, lograron extraer información nutricional de los huesos, lo cual también puede ayudar a discriminar la especie del ser vivo, pudiendo realizar estos estudios incluso en huesos que habían sido sometidos a altas temperaturas [3].

1. Introducción a la espectroscopía

1.1. Desarrollo histórico

Desde la época del Antiguo Egipto, se sabe que el ser humano ya trabajaba en la óptica, siendo capaces de fabricar utensilios como espejos metálicos para aprovechar las propiedades de la luz. En la época de la Antigua Grecia, se intenta dar una explicación al fenómeno de la visión, la escuela Pitagórica, de Euclides y Aristóteles proponen diversas teorías y se logran crear lentes de aumento que permitían observar objetos de forma más detallada [4].

Aunque todas estas teorías son bastante diferentes a lo que conocemos actualmente de la óptica, sirvieron para dar comienzo a una rama de la ciencia en la que aún hoy en día se realizan avances.

Un avance clave en el campo de la óptica lo realizó Isaac Newton en torno a 1665, cuando descubrió que al hacer pasar la luz del Sol a través de un prisma (figura 1), esta se descomponía en los diferentes colores que la componen, formando un espectro. Este experimento le permitió demostrar que la luz blanca está formada por una combinación de todos los colores del espectro visible [5].

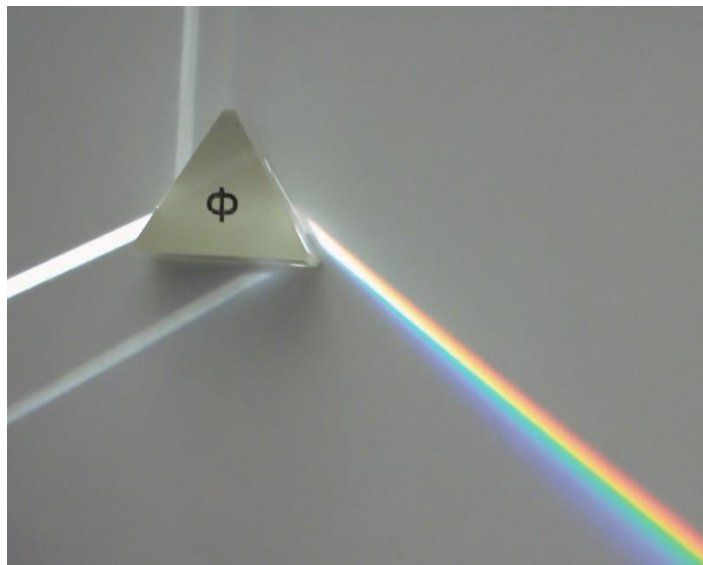


Figura 1. Descomposición de la luz solar en diferentes colores tras pasar por un prisma [6].

Qué es un espectro en este contexto? La RAE lo define como: *distribución de la intensidad de una radiación en función de una magnitud característica* [7]. Centrados en la óptica, esa magnitud puede ser la longitud de onda, la cual puede definirse como la distancia entre dos puntos equivalentes consecutivos de una onda (por ejemplo dos crestas) (Figura 2). Por lo que en este caso, el espectro es la representación de la luz descompuesta en sus diferentes longitudes de onda.

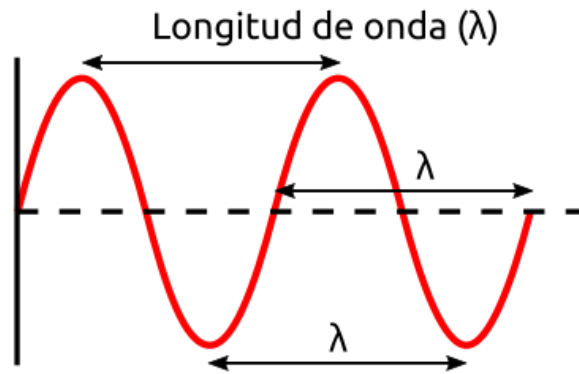


Figura 2. Longitud de onda [8].

Una vez introducido el concepto de espectro, Wollaston construyó uno de los primeros espectrómetros que era capaz de recoger las longitudes de onda emitidas en una radiación de luz, y al visualizar el espectro con la luz emitida por el sol, apreció que se podían vislumbrar unas rayas negras (Figura 3), que parecía que separaban los colores. No fue hasta medio siglo después, cuando Joseph Fraunhofer, Kirchhoff y Bunsen demostraron que esas líneas estaban asociadas con los elementos presentes en la atmósfera del sol [9], siendo así las primeras personas en descubrir que mediante el espectro obtenido de la luz emitida por un objeto se puede estudiar su composición.

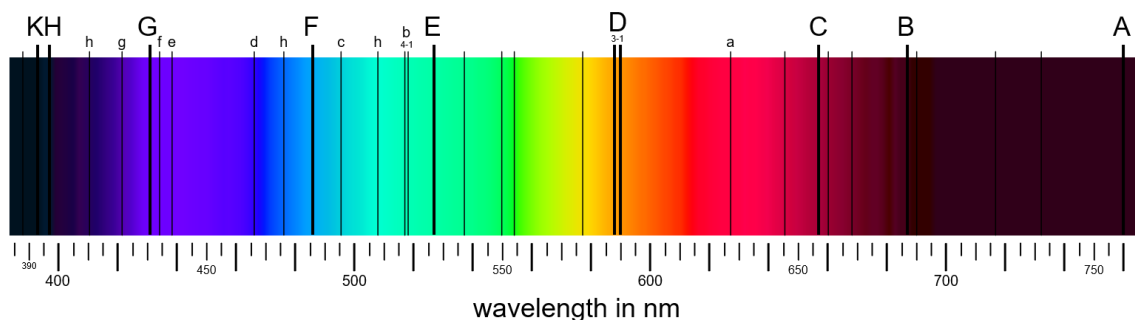


Figura 3. Espectro de la luz solar [10].

De esta forma nace la espectroscopía, definida como la rama de la ciencia que estudia la interacción entre la luz y la materia [11]. Esto es algo muy útil, ya que entendiendo de qué forma emiten luz los objetos en función de los materiales que lo componen, podemos determinar los elementos que lo forman y, sabiendo esto, realizar estudios científicos sobre diversos temas que veremos más adelante.

1.2. Tipos de espectroscopía

Dentro de la espectroscopía podemos encontrar varios tipos, que principalmente se dividen en espectroscopía de absorción y de emisión.

- **Espectroscopía de absorción**

Este tipo de espectroscopía se basa en que cada elemento atómico absorbe ciertas longitudes de onda de la luz que le llega. Por ello, sabiendo las longitudes de onda asociadas a la absorción de un determinado elemento y viendo el espectro de la luz que ha atravesado un objeto/compuesto, podemos saber si ese elemento está contenido en el objeto, ya que en el espectro veremos que hay luz (con la intensidad determinada por la fuente de luz empleada) en todas las longitudes de onda, excepto en las asociadas a los

elementos que contiene el objeto (que aparecerán atenuadas) [11]. Un ejemplo de este tipo de espectroscopía es el de la luz del sol descompuesta en el prisma explicado en la página anterior.

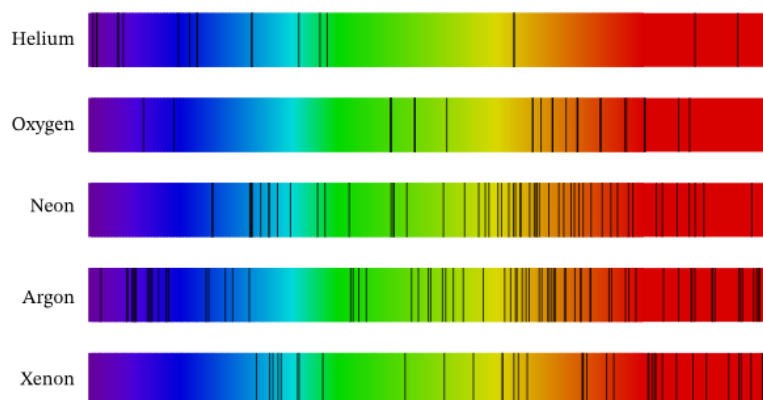


Figura 4. Espectro de absorción de algunos gases [12].

- **Espectroscopía de emisión**

Este es otro tipo de espectroscopía que se produce de forma bastante similar al anterior, diferenciándose en que en la de emisión cada elemento tiene unas "líneas de emisión" que emite a determinadas longitudes de onda asociadas, de nuevo, a determinados elementos atómicos, por lo que en este caso únicamente vemos en el espectro esas líneas de emisión asociadas a cada elemento. Este tipo de espectroscopía requiere aportar energía suficiente a la muestra objeto de estudio como para conseguir que se emita luz durante el proceso. Por ejemplo, con la técnica LIBS, un láser muy intenso emite un pulso de luz que impacta en la superficie de la muestra, ablacionando, es decir, eliminando una parte del material y emitiendo luz en el proceso [13].

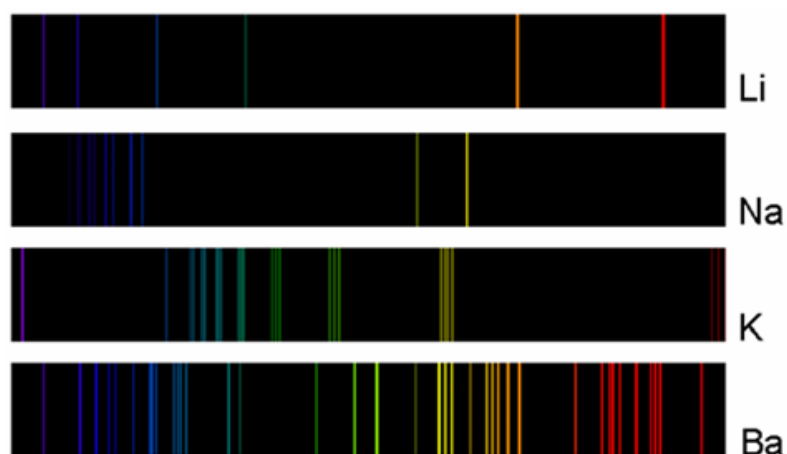


Figura 5. Espectro de emisión de algunos elementos [13].

Estas líneas que aparecen tanto en la espectroscopía de absorción como en la de emisión están asociadas cada una a un único elemento. En el caso de la emisión, aparecen cuando un electrón pasa de un nivel de energía a otro más bajo (Figura 6), ya que el fotón pierde energía que se transforma en luz, y esa energía (energía de *gap*) está inversamente relacionada con la longitud de onda a la que emite el fotón (es lo que se conoce como proceso de emisión espontánea).

Se ha introducido el término energía de gap, definido como la energía mínima necesaria para que un electrón pueda pasar desde la banda de valencia a la banda de conducción. La capacidad de conducción eléctrica de un material dependerá de la cantidad de electrones que posean una energía superior a la energía de gap.

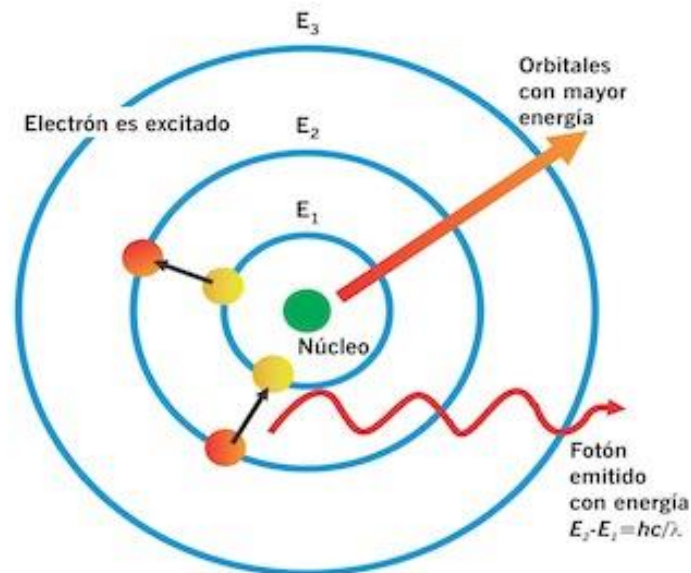


Figura 6. Emisión de luz [14].

En cambio, en el caso de la absorción ocurre lo contrario, el electrón absorbe la energía que le llega en forma de luz, permitiéndole pasar a un nivel de energía superior (un buen ejemplo sería el efecto fotoeléctrico) (Figura 7) .

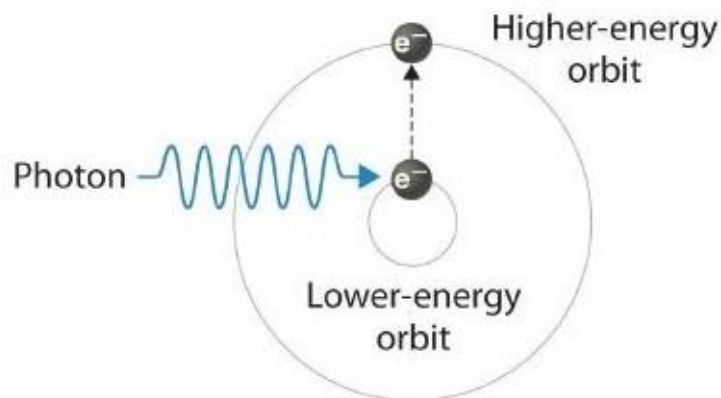


Figura 7. Absorción de luz [15].

Podemos ver la relación entre longitud de onda y energía de *gap* en la siguiente expresión [16].

$$(1) \quad E_g = \frac{h \cdot c}{\lambda}$$

Esta ecuación se obtiene al despejar la relación de Planck-Einstein y la relación entre frecuencia y longitud de onda.

(2)
$$E_g = h * \nu$$

(3)
$$\nu = \frac{c}{\lambda}$$

Siendo h la constante de Planck ($6,63 * 10^{-34}$ J*s), c la velocidad de la luz en el vacío ($3 * 10^8$ m/s), λ la longitud de onda de la luz emitida, E_g la energía de *gap* y ν la frecuencia de la onda electromagnética.

Esto provoca que al tener cada elemento una energía de *gap* característica también tenga una longitud de onda de absorción/emisión propia de dicho elemento, por lo que esta es la forma en la que fijándonos en el color o, siendo más precisos, en la longitud de onda de la luz emitida por un objeto podemos estudiar los elementos que lo componen.

2. Espectroscopía de ruptura inducida por láser (LIBS)

Dentro de la espectroscopía de emisión, la cual puede producirse de forma “natural” (por ejemplo, aprovechando la radiación que se produce en la soldadura) o de forma estimulada, se encuentra la inducida por láser, y que como su propio nombre indica, se produce utilizando una fuente de luz (un láser) para conseguir estimular suficientemente la muestra como para que ésta llegue a emitir luz, generalmente implicando un proceso de ablación. Este tipo de espectroscopía va a ser el que se utilice para el análisis de muestras en este TFG.

2.1. Funcionamiento de los sistemas LIBS

El funcionamiento de esta técnica se basa en el uso de un láser pulsado de muy alta energía con el que se apunta a la muestra que se quiere analizar. De esta forma, mediante el uso de un sistema óptico, se focaliza toda la energía en un pequeño punto en el cual se alcanzan temperaturas muy elevadas (debido a la ionización de los átomos), formándose un plasma que emite una radiación electromagnética continua y eliminándose (ablación) una pequeña parte del material (unas micras). Después, se recoge la luz emitida y se la analiza mediante un espectrómetro, obteniendo el espectro de emisión de la muestra y permitiéndonos conocer los elementos que componen la muestra mediante el análisis de dicho espectro [11][17].

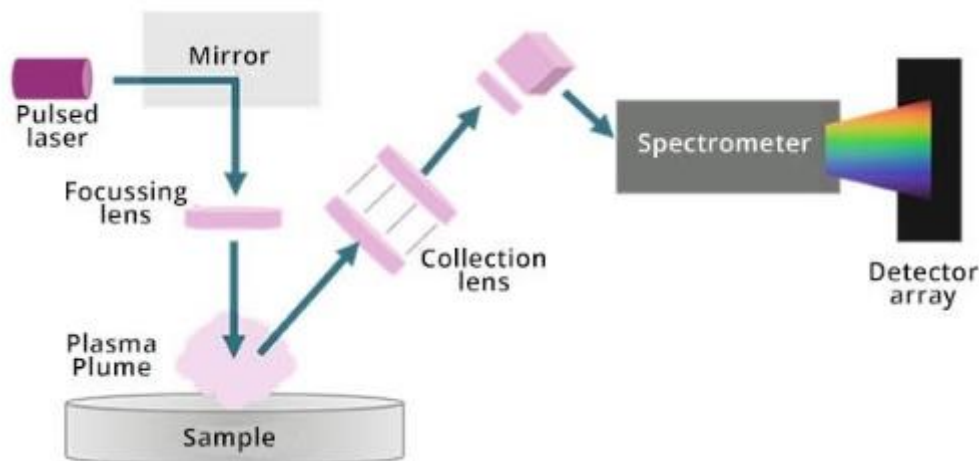


Figura 8. Espectroscopía de ruptura inducida por láser [18].

El montaje de un sistema LIBS consiste en un láser normalmente de tipo Nd:YAG, que emite a una longitud de onda de 1064 nm; aunque también puede emitir a 532 nm, 355 nm o 266 nm; una cámara, que sirve para visualizar la zona de muestreo y colocar de forma óptima la muestra antes de analizarla con el láser; una fibra óptica a través de la cual se transmite la luz recibida hasta el espectrómetro; un espectrómetro en el que se separa la luz que llega del plasma en sus diferentes componentes (longitudes de onda); y un ordenador con un software en el que se controla todo el proceso y dispositivos implicados y se procesa la información [19].

El láser Nd:YAG es un láser que utiliza un cristal de granate de itrio (Y) y aluminio (Al) (YAG), dopado con neodimio (Nd) para la emisión del láser. Puede ser de onda continua o de onda pulsada, el de onda continua se utiliza para aplicaciones industriales (por ejemplo de corte o soldadura) o en medicina; mientras que el de onda pulsada se utiliza en técnicas como

LIBS. Estos láseres proporcionan mucha energía, que puede llegar a ser hasta de Julios (J) en cada pulso, en pulsos cortos de nanosegundos o picosegundos [19][20].

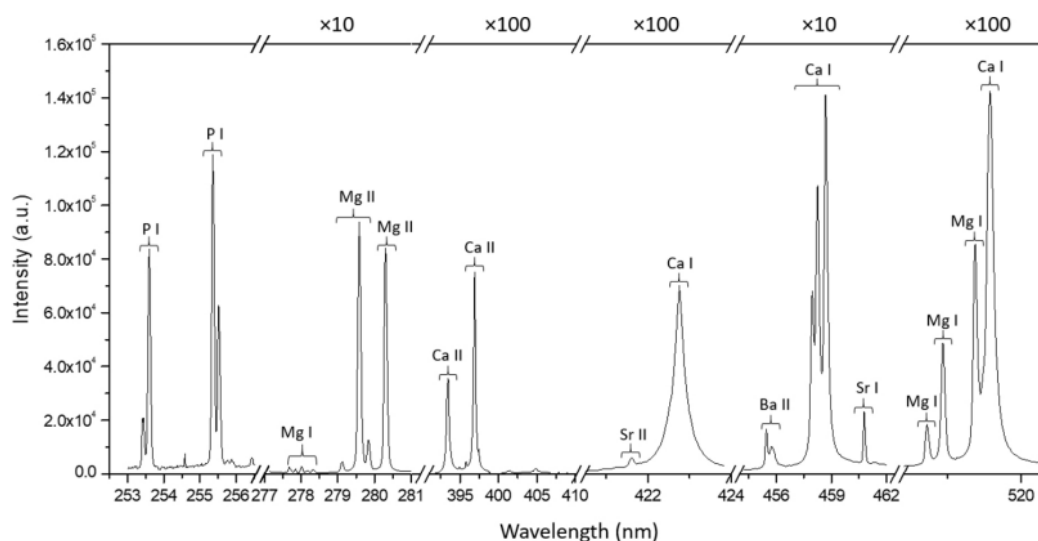


Figura 9. Espectro de emisión para una muestra de hueso [21].

En la figura 9 se proporciona un ejemplo de espectro de emisión, obtenido en un estudio de la composición de varias muestras de huesos. Para la muestra analizada en este ejemplo concreto se observa que contiene elementos como fósforo (P), magnesio (Mg), calcio (Ca), estroncio (Sr) y bario (Ba), cada uno de estos elementos están asociados a líneas de emisión que aparecen a unas longitudes de onda determinadas. Existen bases de datos espectroscópicas con información de estas líneas de emisión, pudiendo así determinar la composición elemental a partir de las longitudes de onda que aparecen con mayor intensidad en el espectro de emisión recogido [22].

Al ser un método que realiza el análisis directamente sobre la muestra de forma rápida y eficaz, en muchos casos se realizan sistemas LIBS, a los cuáles se les incorporan sistemas de autosuficiencia energética, de forma que se obtiene un analizador LIBS portátil, un ejemplo de ello es el sistema LIBS que vemos en la siguiente imagen, creado para el análisis de pesticidas en plantas [23].

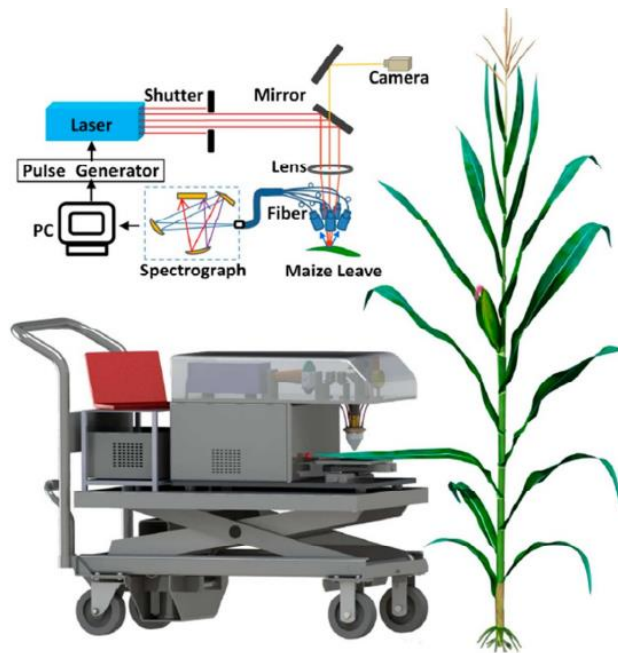


Figura 10. Sistema LIBS portátil para el análisis de pesticidas en plantas [23].

2.2. Ejemplos de estudios realizados con LIBS

Esta técnica sirve para una gran variedad de investigaciones, ya que aporta mucha información sobre la composición elemental de la muestra, sin apenas dañarla y con un sistema no demasiado complejo ni costoso, fácil de transportar y que analiza la muestra rápidamente y en tiempo real. Además de los trabajos de investigación aportados en el contexto de este TFG, otros ejemplos de investigaciones en las que se utiliza la técnica LIBS son:

- **Tesis doctoral de Francisco Anabitarte García: Desarrollo de nuevos dispositivos y técnicas para espectroscopía inducida por láser [17].**

En este trabajo se lleva a cabo un análisis paleonutricional en el que, tras analizar varios huesos de personas, se establece una correlación con las dietas que realizaban dichas personas, clasificándoles en diferentes clases sociales y se trata con ello de averiguar si la economía de las personas influyó en la forma en que les afectó la peste negra. Se realizan las medidas mediante tanto mediante la técnica ICP-OES como con LIBS, de forma que una vez obtenidos ambos resultados se comparan, observando la exactitud de LIBS [24].

La técnica ICP-OES, consiste en un plasma de argón (Ar), que excita los átomos de la muestra, este plasma proporciona energía a los átomos de la muestra, provocándoles elevarse a niveles de energía superiores, cuando los átomos regresan a sus estados originales, se genera emisión de luz, la cual se recoge y procesando dicho espectro de emisión se logra obtener la composición elemental [25].

LIBS y ICP-OES, se diferencian principalmente en que, mientras que en LIBS no es necesaria preparación en la muestra, en ICP-OES sí que hay que realizar una preparación, ya que es necesario preparar una solución líquida. Además, como se verá a continuación, la técnica ICP-OES es más exacta para el análisis de elementos.

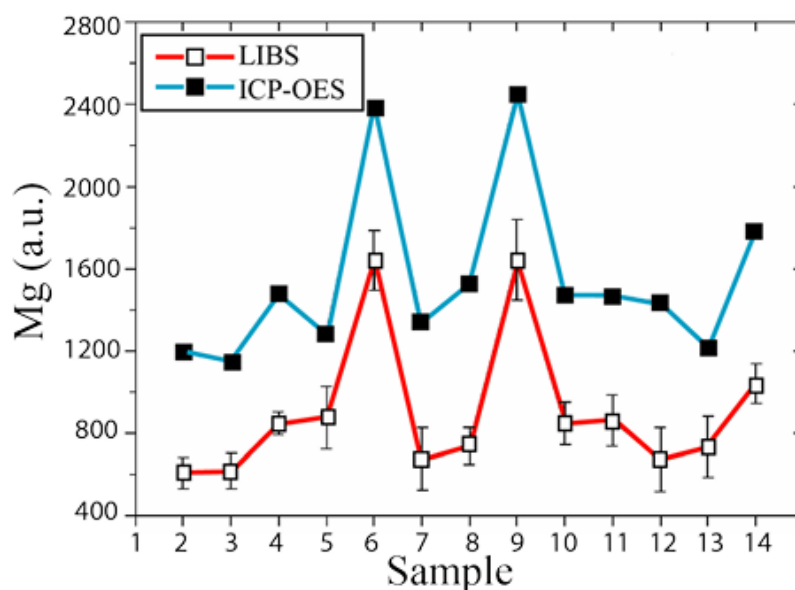


Figura 11. Comparación entre ICP-OES y LIBS para Mg [24].

Se puede apreciar que, aunque los resultados a los que se llega mediante ambos métodos discrepan, hay una tendencia que indica que los valores alcanzados con LIBS son menores que los de ICP-OES, por lo que, si se quiere obtener un valor exacto de una única muestra, la técnica no es totalmente fiable, aunque sabiendo la tendencia que tienen los resultados utilizando LIBS, y procesando los resultados, se podría hacer que estos fueran más exactos. Por otra parte, si se quiere observar patrones o la relación entre varias muestras, LIBS muestra esto de una forma muy precisa, ya que el patrón es prácticamente el mismo en ambos métodos.

Otra investigación que se realiza en esta tesis tiene que ver con la composición de monedas de la época de la República de Roma. De nuevo se analiza la composición mediante LIBS y se comparan los resultados con otros obtenidos anteriormente mediante la técnica XRF (*X-Ray Fluorescence*), de esta forma se vuelve a comparar LIBS con otra técnica, comprobando su exactitud. Además, al ser monedas únicas, las cuales se quiere conservar en el mejor estado posible, esta técnica es perfecta para analizarlas, ya que apenas las daña. También se analiza la composición para averiguar la aleación de las monedas, tratando de comprender las técnicas utilizadas en su fabricación, su ubicación de origen y fecha de muestra [26].

La técnica XRF consiste en un haz de rayos X que excita los átomos de la muestra, provocando que se emita luz, la cual se analiza en un espectrómetro de igual forma que se hacía con LIBS. Por lo tanto, LIBS y XRF son dos técnicas que funcionan de forma similar, ya que no hace falta preparación de la muestra, apenas son destructivas y utilizan un equipo que se puede transportar con facilidad [27].



Figura 12. Monedas romanas analizadas[17].

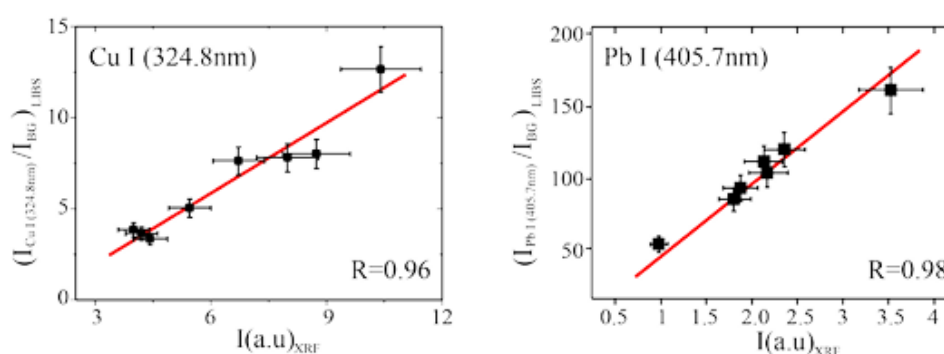


Figura 13. Correlación entre las medidas mediante LIBS y XRF [17].

De nuevo se observa como en el estudio anterior que, aunque LIBS no es demasiado exacto para valores puntuales, sí que indica de una forma precisa los patrones que siguen las muestras.

- **Investigación científica de Jesús Mirapeix, Rosa Arniz-Mateos, Asier García-Escárzaga, Igor Gutiérrez-Zugasti, José Miguel López-Higuera y Adolfo Cobo [28] y Trabajo de Fin de Grado de Ramón Díaz Noval [29].**

Estos son dos proyectos que investigan la composición elemental y sus implicaciones paleoclimáticas en conchas de lapas mediante la técnica LIBS. En ellos, se recolectaron conchas de lapas marinas, analizando la composición de Mg/Ca a lo largo del eje de crecimiento de la concha. Esta técnica tiene el potencial de permitir determinar cuántos años vivió la lapa, en qué estación del año fue recolectada y realizar un estudio sobre la temperatura del mar. Este análisis puede aportar información relevante sobre la evolución del clima en épocas pasadas, así como sobre los hábitos alimenticios y comportamientos de nuestros antepasados.

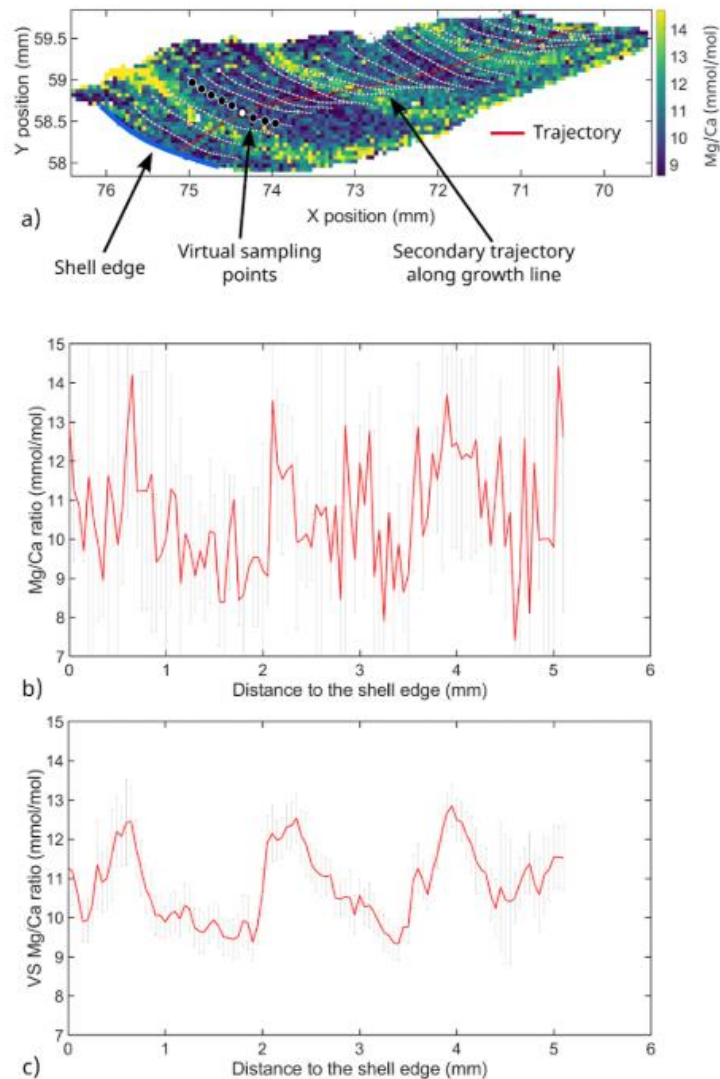


Figura 14. Concentración de Mg/Ca analizada con LIBS [28].

En la figura 14 se observa el resultado del análisis con LIBS, se seleccionó una trayectoria en la concha y se analizó la concentración de Mg/Ca en cada punto de esa trayectoria, obteniendo los resultados que se ven en las figuras b y c.

- **LIBS y CF-LIBS para el análisis de concentración de contaminantes ambientales [30].**

En este artículo, se presenta un estudio sobre la detección y cuantificación de metales pesados tóxicos en muestras ambientales. Para ello se utilizaron las técnicas LIBS y CF-LIBS, las cuales se diferencian en que, mientras que LIBS requiere el uso de muestras de referencia para la calibración, CF-LIBS utiliza modelos físicos y matemáticos para estimar las concentraciones directamente del espectro obtenido [31,32]. Además, se utilizó ICP-OES para comparar los resultados obtenidos.

El estudio analiza las concentraciones de metales pesados en muestras de suelo y se enfatiza la necesidad de evaluar la contaminación en zonas cercanas a industrias. En particular, se centra en el Área de Jajmau, en Kanpur, al ser una zona agrícola a la cual afectan las aguas residuales industriales. Tras la recolección y el análisis de varias muestras de suelo mediante las tres técnicas mencionadas, se encontraron altas

concentraciones de metales pesados, principalmente cromo, cobre, plomo, estroncio, zinc y cobalto.

Finalmente, se comparan los resultados obtenidos con cada técnica, validando tanto LIBS como CF-LIBS como herramientas muy potentes para el análisis espectroscópico.

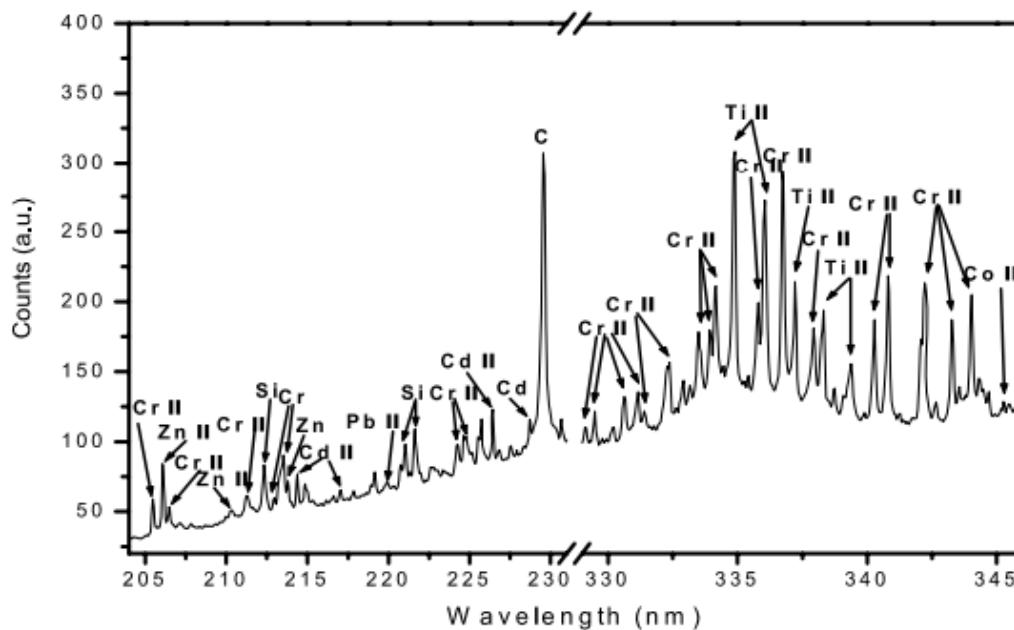


Figura 15. Espectro recogido tras el análisis de una muestra de suelo [30].

2.3. Comparación de LIBS con técnicas espectroscópicas alternativas

Como se ha mencionado anteriormente y se ha observado en los estudios mostrados, aunque LIBS no es una técnica tan exacta para el análisis de una única muestra, como puedan serlo ICP-OES o XRF, es muy útil para identificar patrones al analizar varias muestras, permitiendo además examinar todo tipo de ejemplares sin necesidad de preparación previa. Por ejemplo, con ICP-OES es necesario preparar las muestras en estado líquido, provocando su destrucción, mientras que LIBS, es un método apenas dañino con los especímenes analizados ya que destruye una parte inapreciable a simple vista del material.

Por estos motivos, se considera LIBS la técnica más adecuada para este Trabajo de Fin de Grado, ya que no se busca una precisión exacta para cada muestra individual, sino que se identificarán patrones, tratando de averiguar la especie a la que pertenece cada hueso. Además, dado que se pretende conservar las muestras, es fundamental emplear un método que no implique su destrucción.

3. Estudios centrados en el análisis de huesos

En este proyecto de fin de grado, centrado en el análisis de muestras de huesos mediante LIBS, es fundamental revisar investigaciones previas en las que sustentarse. Además de los estudios que aplican técnicas espectroscópicas como LIBS mostrados anteriormente, también se considerarán trabajos más generales sobre el análisis de huesos, incluso aquellos que no empleen necesariamente estos métodos. Esto permitirá terminar de establecer una base sólida antes de iniciar el trabajo de experimentación. A continuación, se presentan algunos de estos estudios:

- **LIBS y redes neuronales para la discriminación de restos óseos [33].**

En este estudio se emplea la metodología LIBS, combinada con redes neuronales, como una alternativa para discriminar individuos a partir de restos óseos. Para ello, se utilizaron huesos recogidos en la Capilla Unitaria de Mánchester, datados entre los siglos XVII y XIX.

Actualmente, existen diversos métodos para analizar ADN antiguo; sin embargo, presentan ciertos inconvenientes. Por ejemplo, las muestras pueden contaminarse con material genético moderno y, con el paso del tiempo, gran parte del ADN contenido en ellas se degrada. Además, estas técnicas suelen ser considerablemente costosas. Como solución a estas limitaciones, se propone el uso de espectroscopía LIBS combinada con redes neuronales. El análisis mediante LIBS resulta relativamente sencillo, y, si se dispone de un número suficiente de muestras, una red neuronal puede entrenarse para generar un modelo capaz de distinguir entre diferentes tipos de muestras óseas. En este trabajo se analizaron siete huesos pertenecientes a cinco individuos, clasificando los elementos detectados en dos grupos: por un lado, los elementos biogénicos (grupo 1), que forman parte de la matriz ósea; y por otro, los elementos diagenéticos (grupo 2), asociados a la contaminación de las muestras por componentes del suelo. A partir de esta clasificación, se desarrollaron dos modelos distintos utilizando redes neuronales.

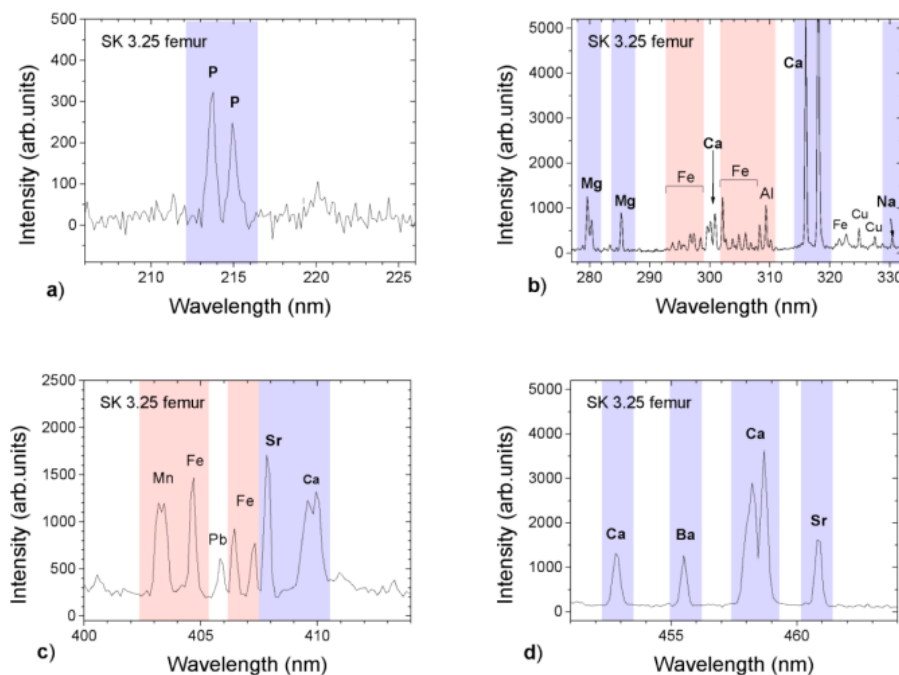


Figura 16. Análisis mediante LIBS de las muestras óseas [33].

Finalmente, ambos modelos fueron utilizados para discriminar entre las muestras. El modelo basado en los elementos biogénicos logró distinguir correctamente los siete huesos con un 100% de efectividad, mientras que el modelo basado en los elementos diagenéticos presentó un rendimiento inferior. Asimismo, se desarrolló un tercer modelo que combinaba ambos enfoques, aunque este no mostró un funcionamiento adecuado.

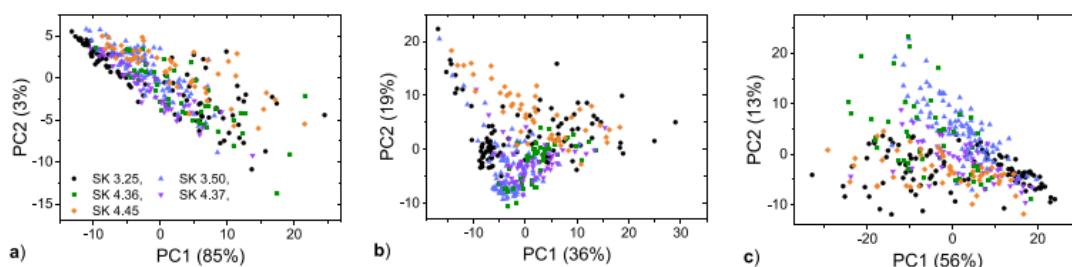


Figura 17. Análisis de los elementos usando el modelo que emplea a) grupo 1, b) grupo 2, c) ambos grupos [33].

- **LIBS para la clasificación e individualización forense precisa de huesos [34].**

Este trabajo se centra en el uso de la técnica LIBS y en el procesamiento de datos como herramientas para la clasificación de individuos a partir de restos óseos. Entre las técnicas de procesamiento de señal empleadas, se destaca la aplicación de algoritmos basados en redes neuronales.

Se presentan varias ventajas de la técnica LIBS: bajo costo de utilización, carácter prácticamente no destructivo, obtención rápida de resultados y la generación de un espectro único por muestra. Esta última característica es especialmente relevante, ya que se ha demostrado que la composición elemental de los huesos es única para cada individuo. Por ello, LIBS se posiciona como una herramienta de gran utilidad en el ámbito forense.

No obstante, diversos factores pueden influir en la composición ósea, tales como la dieta, la edad, el sexo, el estado de salud del individuo y las condiciones del entorno donde se han conservado los restos. Estos factores pueden implicar variaciones en la concentración de elementos, ya que los huesos pueden absorber compuestos presentes en el suelo, como uranio (U), flúor (F) o carbono (C).

Además del análisis espectroscópico mediante LIBS, este estudio pone especial énfasis en el procesamiento de datos. Se utilizaron varios algoritmos de clasificación supervisada basados en redes neuronales. En total, se analizaron 25 muestras óseas procedentes de cinco individuos, evaluando siete métodos de clasificación diferentes. El análisis se centró en las líneas de emisión de calcio, estroncio y magnesio en longitudes de onda específicas.

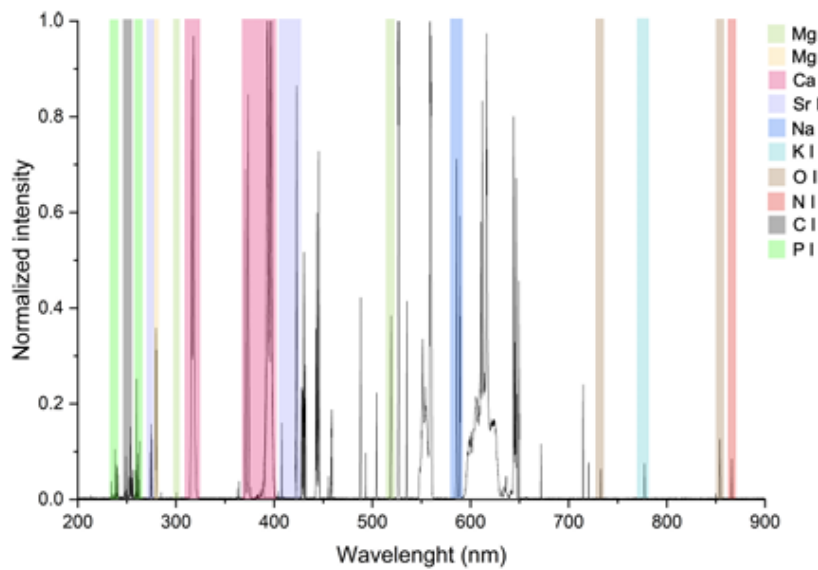


Figura 18. Espectro normalizado obtenido con LIBS para una muestra ejemplo [34].

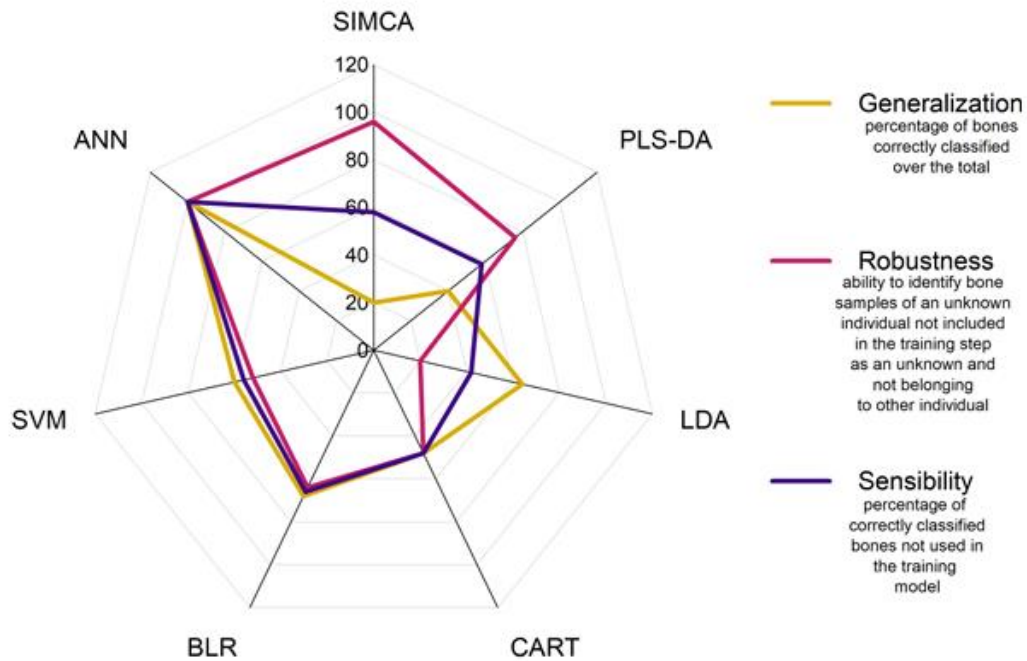


Figura 19. Espectro normalizado obtenido con LIBS para una muestra ejemplo [34].

En la figura anterior se presenta la evaluación de los siete métodos de clasificación empleados, considerando tres métricas clave: la generalización, que representa el porcentaje de muestras óseas clasificadas correctamente por cada método; la robustez, entendida como la capacidad del modelo para identificar correctamente nuevas muestras; y la sensibilidad, que indica el porcentaje de muestras no utilizadas durante el entrenamiento que fueron clasificadas de forma correcta.

En conclusión, este estudio pone de manifiesto las ventajas combinadas del uso de la técnica LIBS y de los algoritmos de redes neuronales para la resolución de problemas forenses complejos.

- **Identificación de fósiles y huesos mediante el uso de LIBS y redes neuronales [35]**

En este estudio se presenta un método que combina LIBS con redes neuronales para el análisis e identificación de restos óseos. El objetivo es abordar la problemática en yacimientos con acumulaciones de restos faunísticos, donde coexisten huesos de numerosos individuos.

Los métodos tradicionales permiten la diferenciación entre especies, pero no entre individuos. Para ello, normalmente se recurre a técnicas genéticas, las cuales son destructivas, costosas y requieren mucho tiempo. En cambio, el uso de LIBS ofrece un análisis rápido, sencillo y mínimamente destructivo, cuya precisión se ve significativamente aumentada al integrarse con redes neuronales, las cuales permiten extraer y procesar información de la composición elemental de las muestras para su clasificación.

En este trabajo se emplearon huesos de ciervos y ovejas alimentados de forma controlada, con el fin de minimizar variaciones externas. Los resultados mostraron una precisión del 100% en los datos de prueba, sin falsos positivos ni negativos.

Asimismo, se analizaron 25 muestras fósiles pertenecientes a cinco individuos con una antigüedad superior al millón de años. A pesar de los procesos *post-mortem* sufridos por los restos, el método permitió una clasificación precisa por individuo. Las muestras que no podían ser clasificadas con suficiente fiabilidad fueron descartadas, lo que permitió mantener una precisión del 100% en los resultados finales.

Por todo ello, se concluye que la combinación de LIBS con redes neuronales constituye una herramienta de gran utilidad para futuros estudios en análisis e identificación de restos óseos.

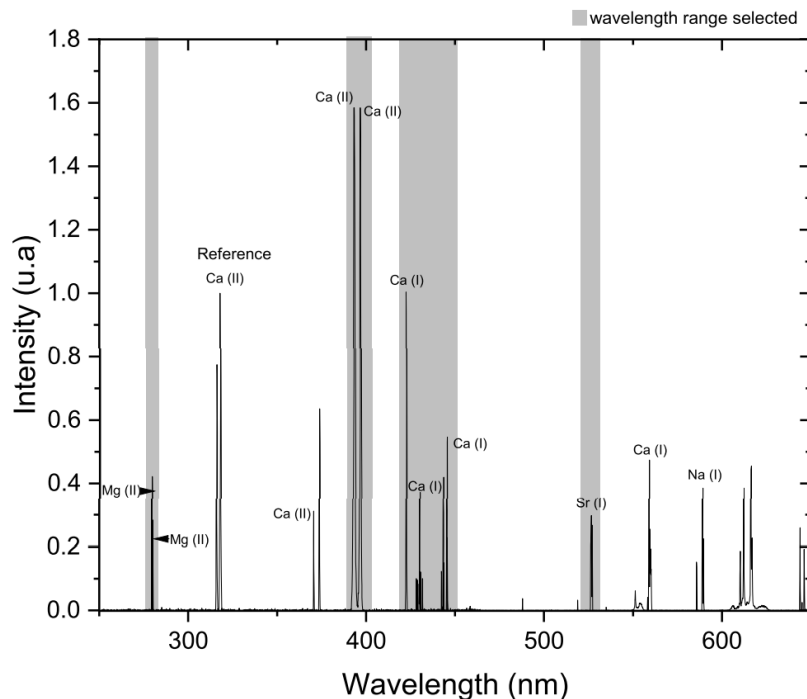


Figura 20. Espectro LIBS típico de una muestra de hueso[35].

- **Identificación humana y análisis de ADN en huesos [36].**

En este trabajo de investigación realizado en 2004, se analizaron huesos arqueológicos para la identificación humana mediante el ADN presente en ellos. Se llegó a la conclusión de que es una forma muy útil para la identificación de individuos, ya que el ADN se conserva durante más tiempo en los huesos en comparación con otros tejidos blandos, con las limitaciones de que el ADN presente en los huesos se degrada con el paso del tiempo, por lo que se dificulta su análisis cuanto más tiempo haya transcurrido desde el fallecimiento del individuo, por otro lado, se comprobó que determinados factores ambientales como la baja humedad y la baja temperatura favorecen su preservación

- **Análisis del canal medular del hueso bovino para la estimación de la edad [37] y Estudio métrico del canal medular del húmero como indicador de la edad [38].**

En estos dos estudios se trata de estimar la edad de individuos bovinos y humanos utilizando sus huesos.

Midiendo la longitud total del hueso y la longitud del canal medular, se establece el índice medular óseo (IMO), el cuál es muy útil para estimar la edad del animal cuando murió. A continuación, se pueden ver en la figura 21 las medidas tomadas en los huesos, siendo LCM la longitud del canal medular y LH la longitud del hueso. Además, en la ecuación 4 se observa el cálculo del IMO.



Figura 21. Medidas tomadas en hueso bovino [37].

(4)
$$IMO = \frac{LCM}{LH} \times 100$$

Por otra parte, para estimar la edad de los individuos humanos se realiza una relación entre edad y nivel del canal medular, obteniendo una regresión lineal que servirá como estimador para la edad en nuevas muestras.

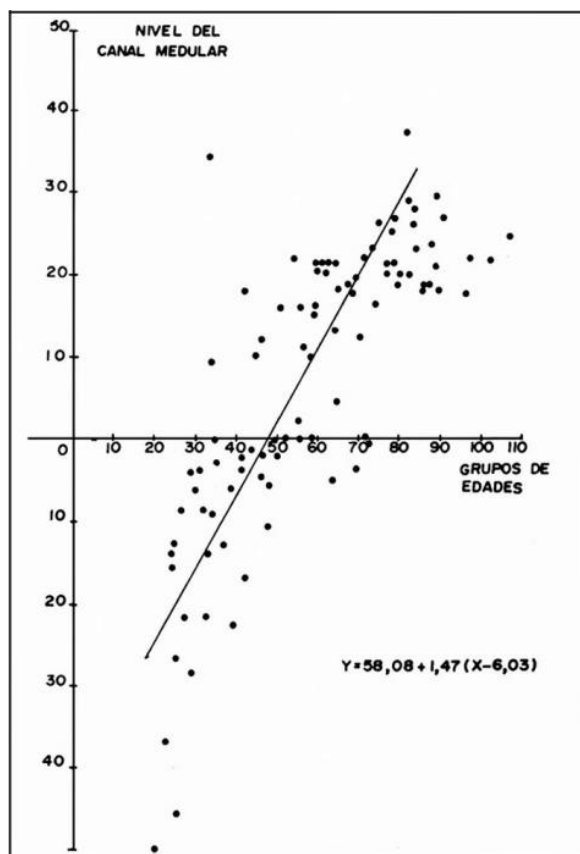


Figura 22. Relación entre nivel del canal medular y edad. Regresión lineal [38].

Se puede observar que, aunque esta regresión lineal no es totalmente exacta, es de gran utilidad para realizar una buena aproximación de la edad de los individuos.

- **Identificación de Especies Mediante el Análisis de Colágeno Óseo por Espectrometría de Masas MALDI-TOF [39].**

En este trabajo se trata de identificar muestras óseas dañadas por el calor o por el tiempo. Para ello se analiza la proteína de colágeno en ellas, ya que es menos susceptible a la contaminación en muestras antiguas.

Para el análisis se utilizó una variante de la técnica ZooMS, que consiste en extraer el colágeno de una muestra ósea, se utiliza una enzima para romperlo en fragmentos y se utiliza espectrometría de masas (MALDI-TOF MS) para analizar la huella resultante y compararla con bases de datos de referencia que permiten identificar la especie.

Se realizó el análisis en 32 especies diferentes de mamíferos, obteniéndose 92 marcadores que se pueden utilizar para la identificación de especies. Además, el método demuestra ser capaz de diferenciar entre especies morfológicamente similares como por ejemplo entre ovejas y cabras.

En las figuras 23 y 24 se pueden observar las especies analizadas en el estudio y los espectros para muestras óseas de conejo y camello.

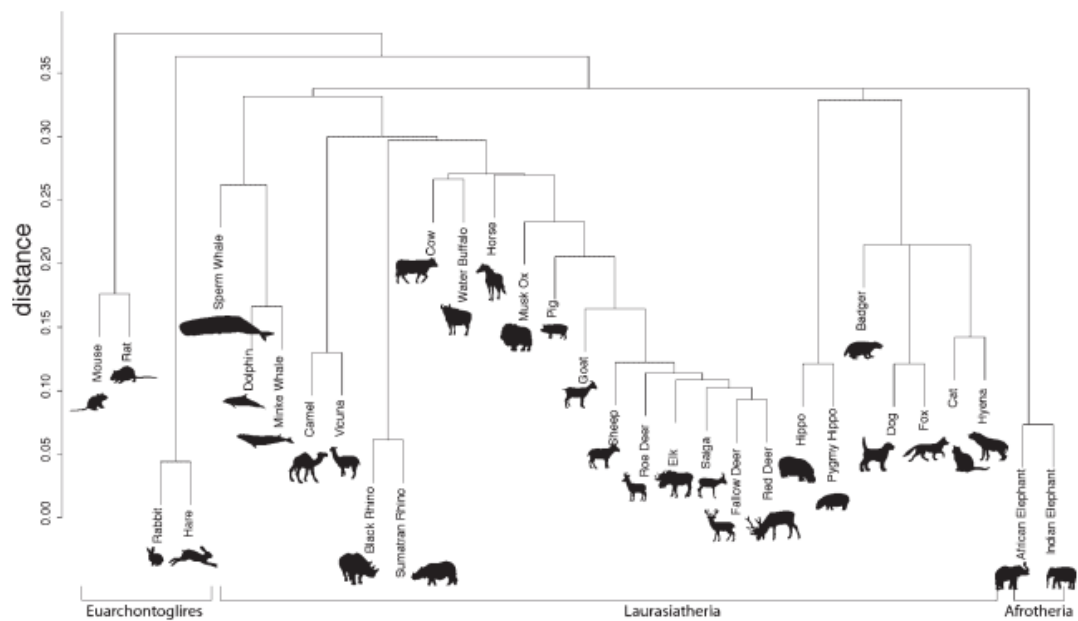


Figura 23. Especies analizadas en el estudio [39].

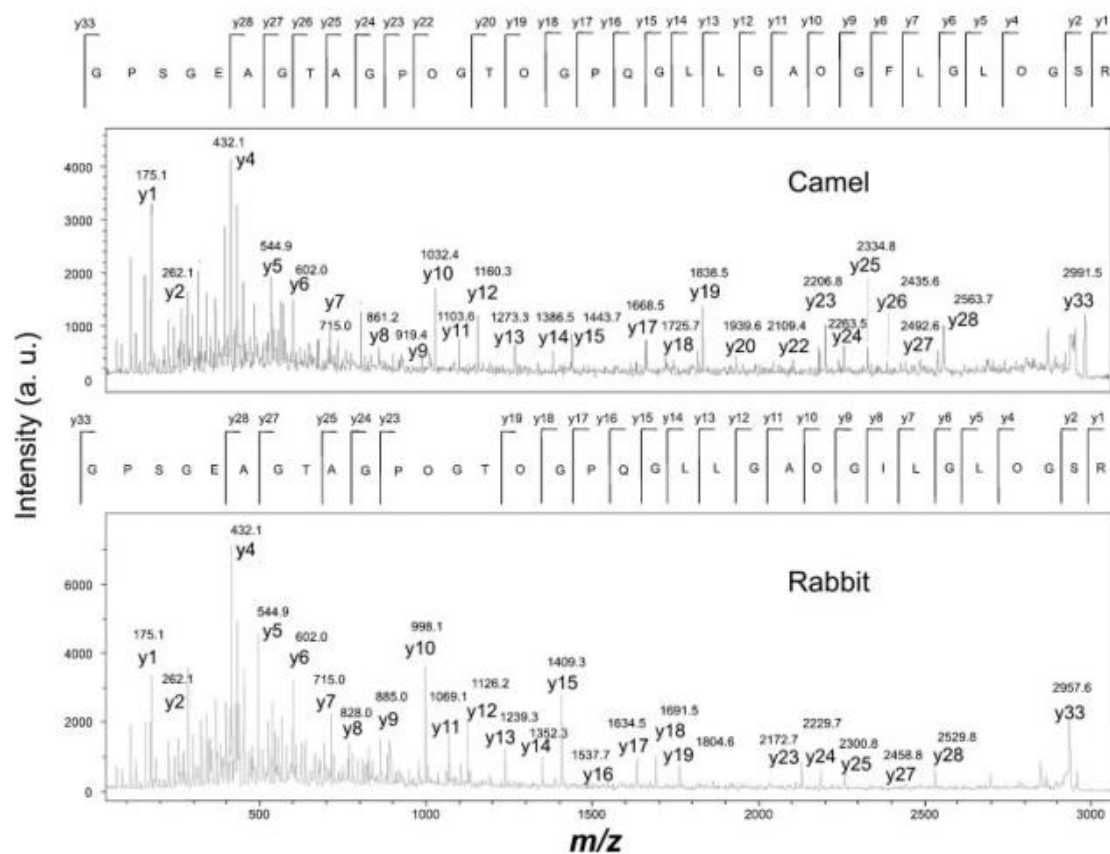


Figura 24. Espectros para muestras óseas de conejo y camello [39].

- **ADN para la Identificación de Huesos Arqueológicos de Animales**

Se realizó una investigación en la que se analiza el ADN para la identificación de especies a partir de fragmentos de huesos arqueológicos.

Se plantea la problemática de los arqueólogos para identificar especies a partir de huesos recuperados, además, se dice que algunas especies son morfológicamente indistinguibles, por lo que no se pueden distinguir sus huesos a simple vista. Para solucionar dichos problemas, se utiliza la metodología PCR/ADN, con la cual, utilizando muestras recogidas en Canadá, datadas en una antigüedad de entre 1000 y 5400 años, desarrollando un algoritmo que se aplica a huesos encontrados en Nigeria y Camerún.

En la figura 25 se observan los huesos que se utilizaron para el análisis, indicando su ubicación de procedencia, el tipo de hueso, la identificación morfológica y la obtenida tras la identificación molecular. En la figura 26 se muestra un intento de identificación de las especies presentes en cada reacción.

Bone ID	Site	Layer	Fragment	Morphological identification	Molecular identification ^a
2562	HSI	Old Women	UID screen	Bison?	Bison
2532	HSI	Old Women	Carpal	Bison	Bison
2550	HSI	Old Women	UID screen	Bison?	Bison
18533	HSI	Avonlea	Humerus	Bison	—
18514	HSI	Avonlea	Vertebra	Bison	—
18536	HSI	Avonlea	Radius	Bison	Bison
76984	HSI	Pelican Lake	Phalanx	Bison	—
76993	HSI	Pelican Lake	Rib	Bison	Bison
76994	HSI	Pelican Lake	Innominate	Bison	Bison
76999	HSI	Pelican Lake	Tibia	Bison	Bison
84046	HSI	Mummy Cave	Rib	Bison	—
84049	HSI	Mummy Cave	UID screen	Bison?	—
84084	HSI	Mummy Cave	Metapodial	Bison	—
1284-2	PMW636	3	L. scapula	Goat	—
1284-1	PMW636	3	L. scapula	Sheep	—
1295-1	PMW636	5	R. metacarpal	Sheep	—
1267-1	PMW636	2 – 20-30cm	Long bone	Cattle	—
1267-2	PMW636	2 –	Cervical	Cattle	Sheep
1235-2	PMW618	7 – 40-50cm	Rib shaft	Unknown	Goat
1235-1	PMW618	7	UID	Unknown	Goat
1351-1	PMW635	7	Long bone	Unknown	Sheep
45-1	Sukur	7	Thoracic vert. spine	Cattle	—
52-1	Sukur	8	Rib shaft	Cattle	—
52-2	Sukur	8	Rib shaft	Cattle	—
7-4	Sukur	2	L. distal metacarpal	Cattle	—
93-1071-1	Sukur	23	Rib	Unknown	Goat
93-475-1	Sukur	10	Carpal/tarsal?	Unknown	—

—=unable to PCR.

Figura 25. Muestras de huesos utilizadas para el estudio [40].

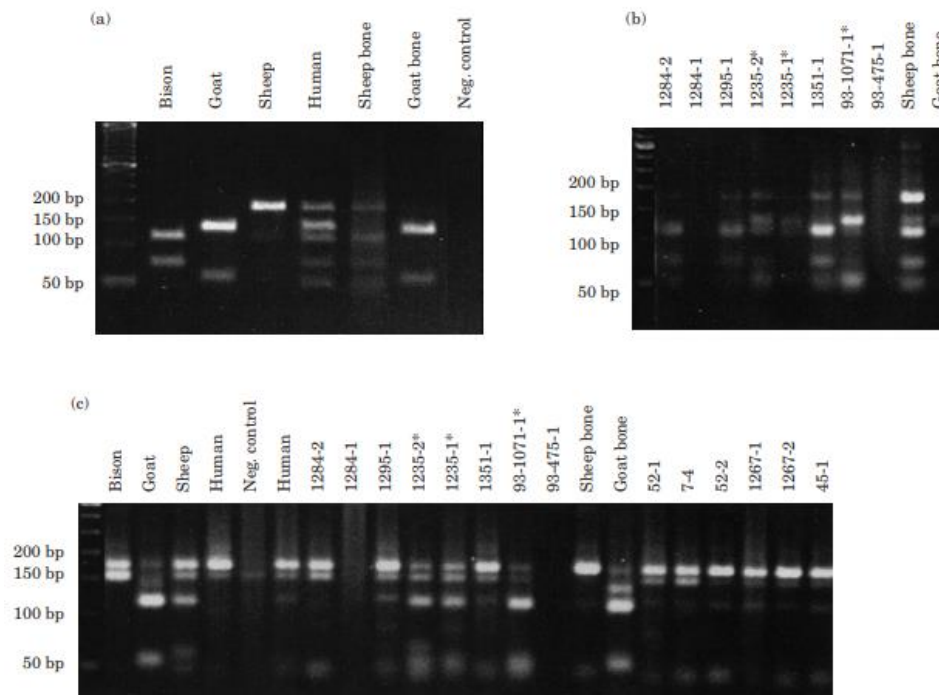


Figura 26. Identificación de especies presentes en cada reacción [40].

Se consiguió distinguir entre huesos de bison, oveja y cabra, resaltándose así la funcionalidad de la técnica, en cambio, se necesita una técnica más avanzada para la diferenciación entre animales de la misma especie.

4. Montaje del sistema LIBS utilizado

A continuación, se puede observar el montaje del sistema LIBS que se ha utilizado para tomar las medidas en las muestras y obtener los datos que se utilizan en este TFG.

En primer lugar, se muestra un esquema general del sistema LIBS.

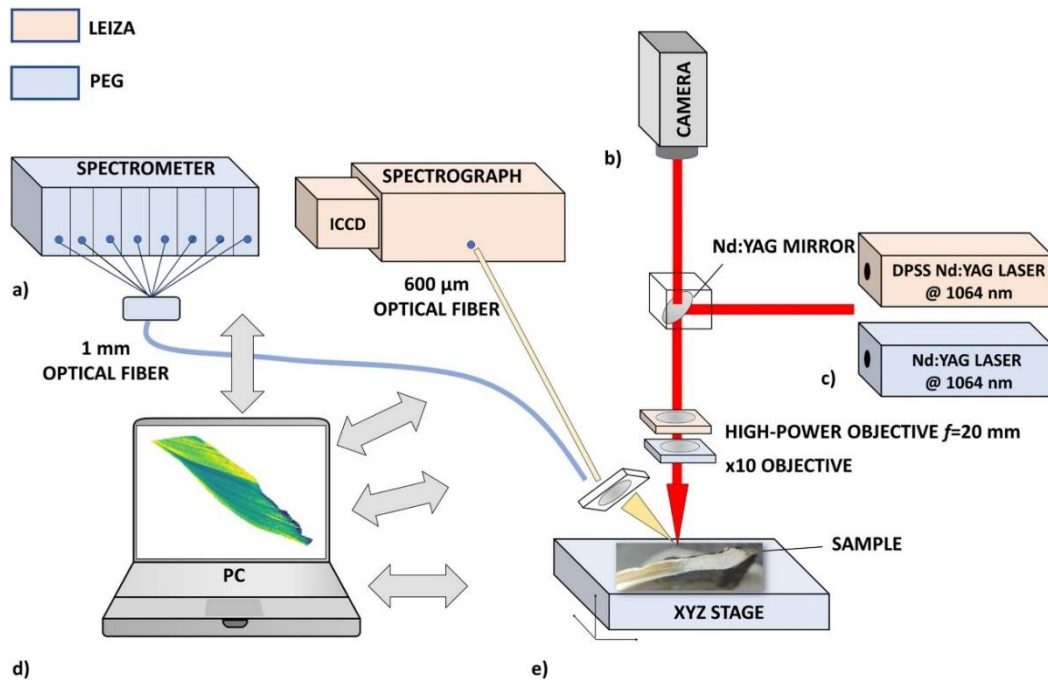


Figura 27. Esquema del sistema LIBS utilizado [41].

El sistema LIBS utilizado para este trabajo está compuesto por un láser pulsado, varios espejos y lentes que llevan y focalizan el haz láser hacia la muestra, un soporte motorizado donde se coloca la muestra, una cámara que permite monitorizar el proceso, una fibra óptica que recoge la luz emitida y la dirige al espectrómetro y un control remoto para seleccionar los parámetros del láser.

El láser utilizado es un Q-Switched Nd:YAG double-pulsed (Lotis LS-2134D), operando en una longitud de onda de 1064 nm con una tasa de 10Hz y utilizando pulsos de 16 ns con una energía de 35 mJ. El modelo de láser empleado se puede observar en la siguiente figura.



Figura 28. Láser utilizado para la experimentación. Elaboración propia.

A continuación del láser, se encuentran varias lentes que sirven para guiar y focalizar el rayo hacia la muestra de forma correcta.

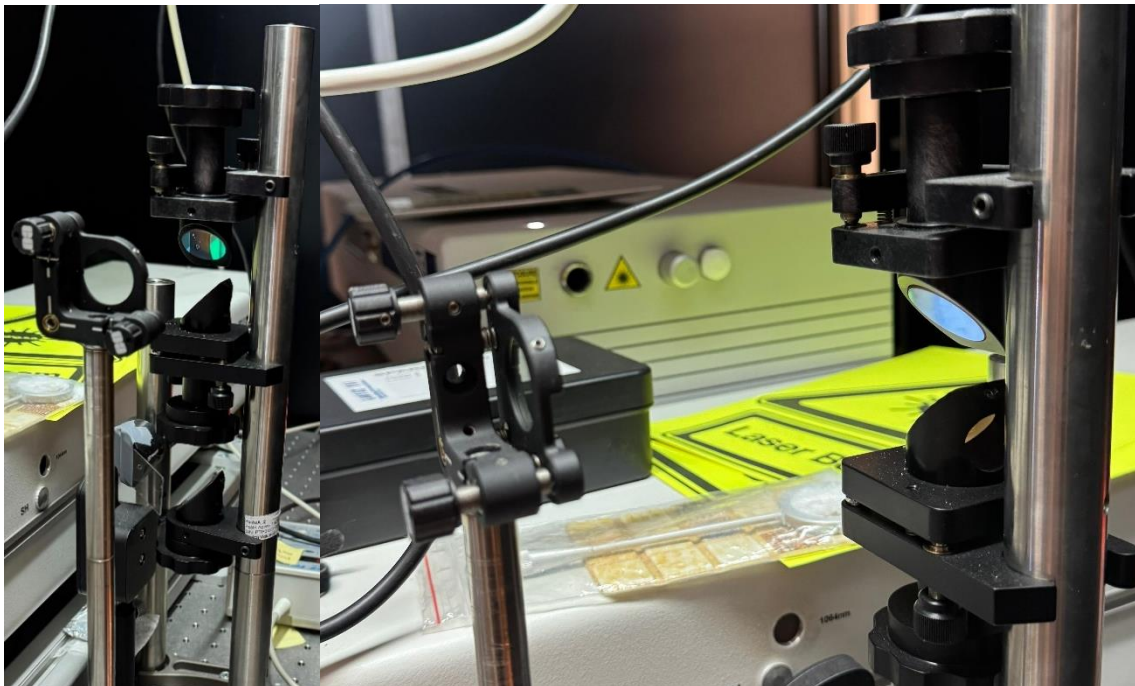


Figura 29. Lentes para la focalización del láser hacia la muestra. Elaboración propia.

Después de las lentes, está situado el soporte donde se coloca la muestra, el cuál esta provisto de un sistema de motores, que se utiliza para mover la muestra hacia el punto exacto en el que se quiere que el láser impacte. Este motor también es útil para realizar

varios disparos del láser en diferentes puntos de forma automatizada, es decir, se le indica al programa cuántos disparos y en qué puntos se quieren realizar y el motor mueve de forma autónoma la muestra siguiendo trayectorias o cubriendo ciertas superficies especificadas anteriormente mediante el programa de control.

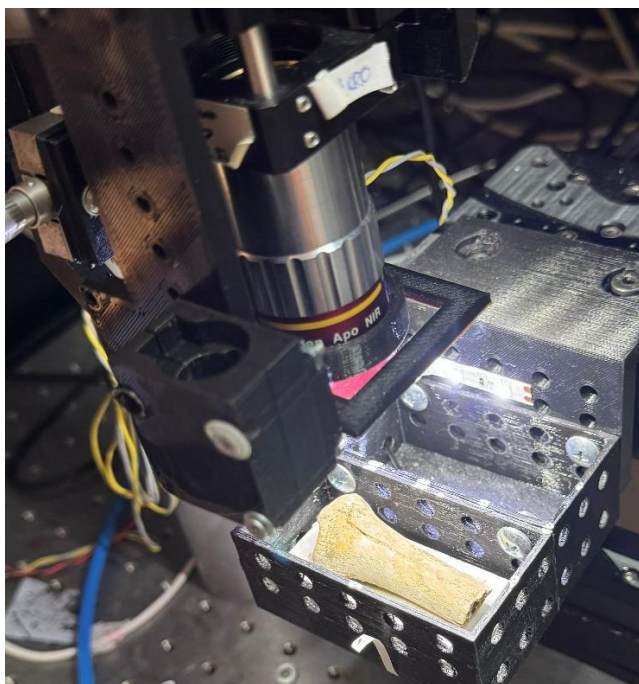


Figura 30. Soporte motorizado. Elaboración propia.

Encima de este soporte hay varias lentes que recogen la luz generada y la llevan mediante un cable de fibra óptica de 1 mm de diámetro hasta un espectrómetro ULS2048-USB2_RM CCD de ocho canales con un rango espectral de 178 a 889 nm y con una resolución de entre 0,015 y 0,06 nm.

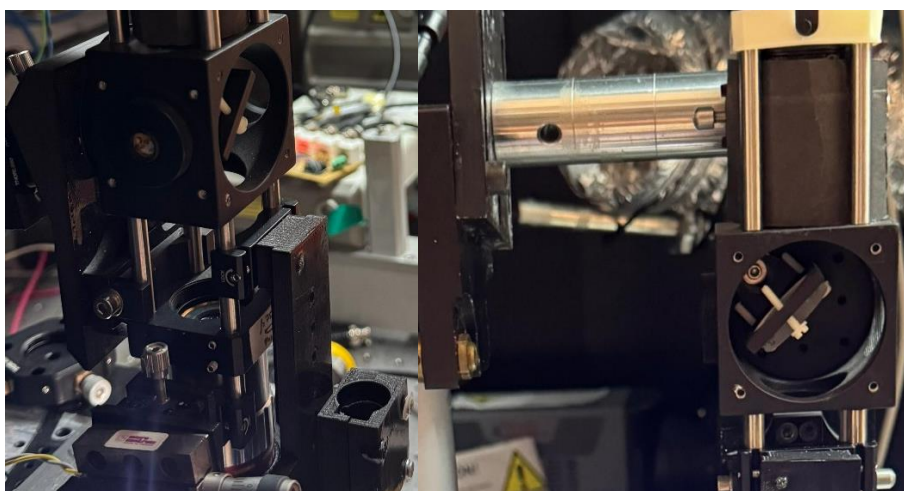


Figura 31. Lentes para la recolección de la luz generada. Elaboración propia.

Por último, se encuentra una cámara, la cual está situada encima del soporte y de las lentes, que se encarga de mostrar al usuario en la pantalla el punto de la muestra que se está analizando, para que el usuario pueda colocar la muestra adecuadamente y dirigir el láser hacia el punto deseado. Además, dicha cámara enfoca con 10 aumentos para que el usuario visualice de una mejor forma la muestra que está analizando.

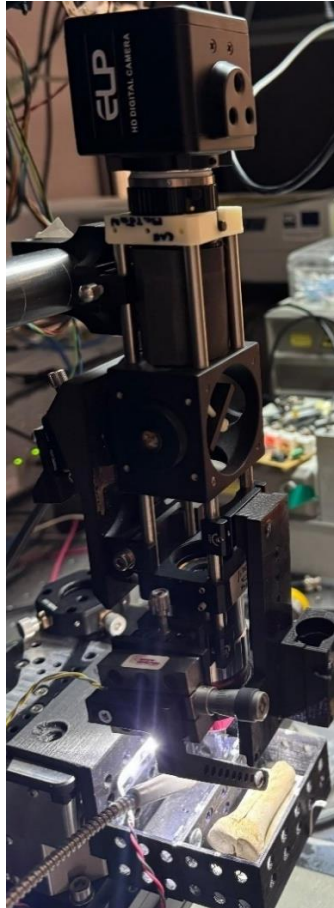


Figura 32. Cámara digital para la visualización de la muestra. Elaboración propia.

Por otra parte, el control remoto permite controlar los parámetros del láser en función del tipo de experimentación que requiera el usuario.

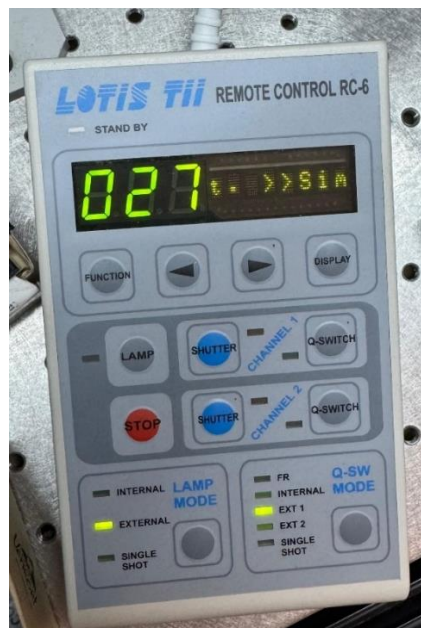


Figura 33. Control remoto. Elaboración propia.

5. Modelos de clasificación y redes neuronales

5.1. Introducción al *machine learning*

El *machine learning* es una rama de la inteligencia artificial, en la cual los sistemas aprenden automáticamente a partir de datos, sin necesidad de ser programados para una tarea en concreto. Estos algoritmos, identifican patrones en los datos de entrenamiento que les son entregados y mejoran su desempeño cuanto mayor número de datos reciben. El objetivo de esta rama consiste en crear algoritmos que sean capaces de realizar predicciones y tomar decisiones de forma autónoma a partir de los datos que han recibido previamente.

Dentro del *machine learning*, existen los algoritmos de regresión y de clasificación. Su diferencia radica en que en la regresión, el modelo predice un valor real cualquiera para cada serie de datos que se le proporcionen, mientras que en la clasificación asigna una de las categorías previamente definidas por el programador en el entrenamiento del modelo. En este trabajo, se van a utilizar únicamente los algoritmos de clasificación.

Respecto al entrenamiento de los algoritmos, existen dos tipos de aprendizajes:

- Aprendizaje supervisado

Durante el entrenamiento, al modelo se le proporciona un conjunto de datos acompañado de sus etiquetas (valores reales). A partir de esta información, el modelo aprende la relación entre las entradas y las salidas, con el objetivo de predecir las etiquetas correspondientes cuando se le presenten nuevos datos que no ha visto previamente.

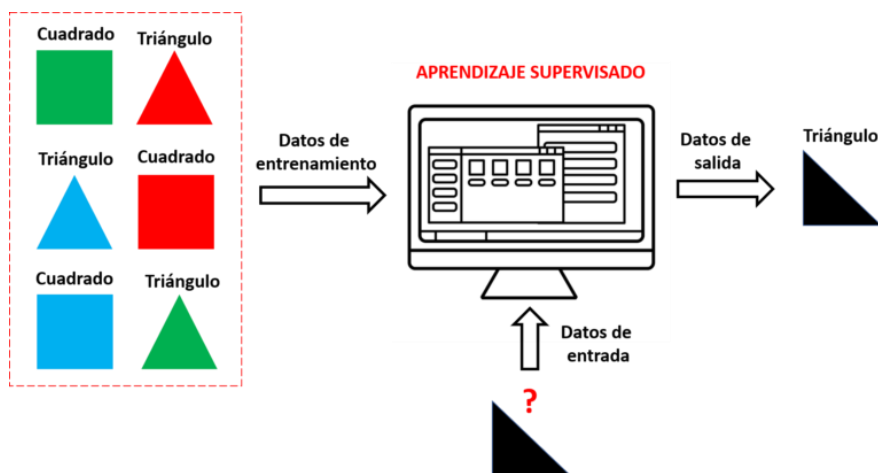


Figura 34. Aprendizaje supervisado [42].

- Aprendizaje no supervisado

En este caso, al modelo se le entrega únicamente un conjunto de datos sin etiquetas. El objetivo consiste en identificar patrones, estructuras o agrupaciones presentes en los datos, sin contar con ejemplos previos de la salida esperada.

En este estudio únicamente se va a utilizar el aprendizaje supervisado, ya que los datos utilizados contienen además las etiquetas de salida.

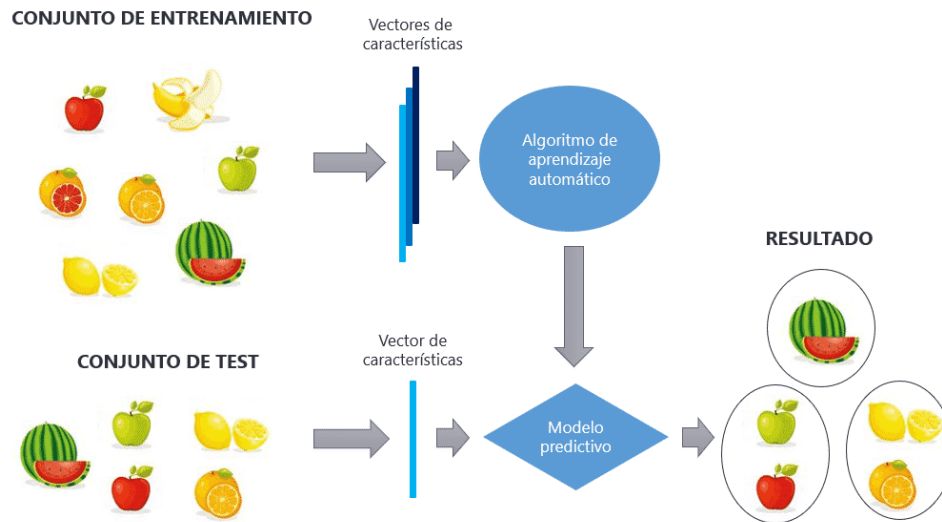


Figura 35. Aprendizaje no supervisado [43].

5.2. Tipos de modelos de clasificación en Matlab

Para este trabajo, además de los espectros obtenidos mediante la técnica LIBS, adquiere una especial relevancia el tratamiento de los datos, dado que de ello dependerá la capacidad del modelo para predecir las especies. Es por ello que, a continuación, se presentarán diversos enfoques de procesamiento de datos ofrecidos por Matlab, que pueden ser de gran utilidad para este estudio.

Matlab incorpora diversas aplicaciones orientadas al desarrollo de modelos predictivos, entre las cuales destacan *Regression Learner* y *Classification Learner*, que como se ha explicado anteriormente, se diferencian en que la primera sirve para algoritmos de regresión, mientras la segunda para modelos de clasificación. Dado que en este trabajo se pretende clasificar espectros en varias especies, se va a utilizar *Classification Learner*.

Con respecto a *Classification Learner*, dentro de dicha aplicación se pueden escoger diversos modelos de clasificación. A pesar de que todos ellos tienen el mismo objetivo, no realizan los cálculos de la misma forma, por lo que sus resultados pueden variar en función de la tarea realizada. Por lo tanto, la elección del modelo resulta fundamental para garantizar la precisión de las predicciones. A continuación se presentan los tipos de modelos de clasificación ofrecidos por Matlab.

- **Árboles de decisión**

Es un modelo con estructura jerárquica en forma de árbol, donde cada nodo representa una pregunta y en función de la respuesta se pasa a un nodo u otro. Algunas de sus ventajas son las siguientes: son muy visuales, es decir, al realizar predicciones se puede ver de forma sencilla porque el modelo ha seleccionado el valor predicho y no otro; funciona tanto con datos numéricos como categóricos (por ejemplo palabras), y es sencillo determinar la fiabilidad del modelo con pruebas estadísticas.

Por otro lado, tiene como principales desventajas que al entrenarles con matrices de datos demasiado grandes, se crean árboles excesivamente complejos, lo cual provoca sobre-

ajuste; además, una pequeña variación en los datos de entrenamiento puede provocar que el árbol cambie completamente; por último, es necesario entrenarles con la misma cantidad de datos para cada categoría, ya que de no ser así, es muy probable que se creen árboles sesgados, los cuales en la mayoría de ocasiones estimen como resultado la categoría que se ha entrenado con una mayor cantidad de datos.

En la figura 36 se observa un ejemplo de un árbol de decisión sencillo.

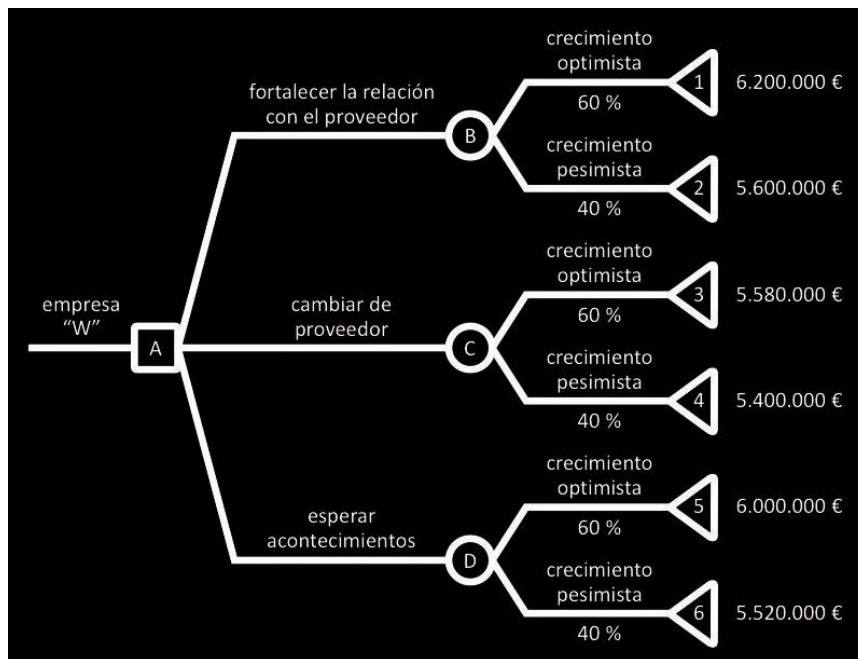


Figura 36. Ejemplo de árbol de decisión [44].

Dentro de los árboles de decisión, Matlab permite escoger entre los tipos *Coarse Tree*, *Medium Tree*, *Fine Tree* y *Optimizable Tree*. Se diferencian en el tamaño del árbol, de menor a mayor respectivamente, excepto *Optimizable Tree*, que consiste en que utilizando optimización Bayesiana, se buscan automáticamente los mejores parámetros en función de los datos de entrenamiento, sin embargo, tiene la problemática de que su coste computacional es muy elevado y por tanto el entrenamiento será demasiado lento en caso de tener muchos datos.

- **Análisis discriminante**

Es un clasificador que trata cada variable como una variable aleatoria gaussiana, estima su media y varianza y usa las estimaciones para calcular la probabilidad de que las nuevas series de datos pertenezcan a cada categoría, asignándoles la categoría con mayor probabilidad. Es decir crea un mapa donde dibuja unas fronteras, colocando los nuevos puntos en el mapa y decidiendo su categoría dependiendo de las fronteras a las que pertenezca.

Dentro de Matlab existen los tipos *Linear discriminant*, *Quadratic discriminant* y *Optimizable discriminant*. Se diferencian en que *Linear discriminant*, crea fronteras de decisión lineales y supone que todas las clases tienen la misma matriz de covarianza, por lo que este modelo funciona correctamente cuando las clases tienen varianzas similares. Mientras que *Quadratic discriminant*, utiliza fronteras de decisión curvas y usa una covarianza diferente para cada clase, este modo se ajusta mejor a los datos pero tiene

mayor coste computacional y es más probable que provoque sobre-ajuste que el modelo lineal. Por otro lado, el modo optimizable, igual que ocurría con los árboles de decisión, ajusta los parámetros del modelo automáticamente en función de los datos de entrenamiento [45,46].

- **Regresión logística**

La regresión logística, es un método de clasificación utilizado normalmente en casos con únicamente dos categorías. Al igual que el resto de métodos de clasificación, calcula la probabilidad de pertenecer a una categoría. Para ello calcula una combinación lineal de las variables y le aplica la función logística (ecuación 5), el resultado obtenido de ésta es una probabilidad, en caso de ser mayor que 0,5 se considerará que pertenece a la clase 1, mientras que si es menor de 0,5 pertenecerá a la clase 2 [47, 48]. Generalmente, cuando se utiliza la regresión logística para casos con más de dos categorías se utiliza la regresión logística softmax, pero esta no está directamente implementada en Matlab, sino que existe la regresión logística eficiente.

$$(5) \quad g(z) = \frac{1}{1 + e^{-z}}$$

Además del modo mencionado, en Matlab también existe el modo *binary GLM logistic regression*, pero la aplicación no permite usarlo en este caso, ya que es solo para clasificaciones con dos categorías. Por lo que se usará únicamente el modo *efficient logistic regression*, que utiliza la función de Matlab `fitcecoc` en caso de variables con más de dos categorías, y la cual combina clasificadores binarios con ECOC (*Error-Correcting Output Codes*), que convierte un problema de n clases en varios problemas binarios y combina los resultados obtenidos para obtener los resultados finales [49].

- **Naive Bayes**

Es un algoritmo que asume que las variables son independientes entre sí y utiliza el Teorema de Bayes [50] para calcular las probabilidades a posteriori, asignando la clase con mayor probabilidad. A continuación se puede observar en la ecuación 6 el Teorema de Bayes, siendo $P(c)$ la probabilidad a priori de la clase, $P(x|c)$ la verosimilitud, $P(x)$ la probabilidad a priori del vector de características y $P(c|x)$ la probabilidad a posteriori de la clase.

$$(6) \quad P(c|x) = \frac{P(x|c)P(c)}{P(x)}$$

Hay tres tipos de clasificadores de *Naive Bayes* en Matlab: *Gaussian Naive Bayes*, *Kernel Naive Bayes* y *Optimizable Naive Bayes*. Los dos primeros se diferencian en que mientras que el de tipo Gaussiano trata las características como variables con distribución normal, el de tipo Kernel estima la distribución de cada variable, por lo que se adapta mejor a los datos y es más exacto en sus estimaciones, pero tiene un mayor coste computacional. Por otro lado, el tipo optimizable, como ocurre siempre en Matlab, adapta los parámetros a los datos de entrenamiento automáticamente.

- **Clasificadores de Vectores de Soporte (SVM)**

Busca un hiperplano que sea capaz de separar la mayor distancia posible los puntos de una clase de los de otra, de forma que se establecen fronteras y al introducir nuevos puntos, estos se sitúan en una de las zonas, de forma que se le clasifica en la clase en la

que esté situado. Este clasificador no necesita almacenar todos los datos de entrenamiento, sino que guarda los subconjuntos de puntos más cercanos al hiperplano separador (vectores de soporte), de forma que con esos vectores es capaz de calcular el hiperplano [51].

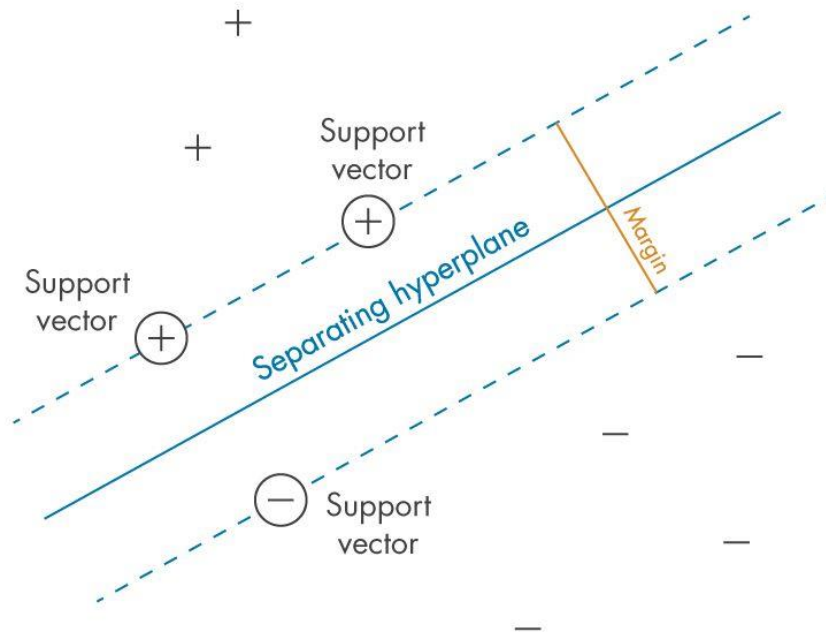


Figura 37. Hiperplano separador y vectores de soporte [51].

En Matlab existen los tipos *Linear SVM*, *Quadratic SVM*, *Cubic SVM*, los cuales se diferencian en que buscan un polinomio lineal, de grado 2 y de grado 3 respectivamente. Por otro lado, también existen *Fine Gaussian*, *Medium Gaussian* y *Coarse Gaussian*, que utilizan fronteras de decisión no lineales. *Fine Gaussian* utiliza fronteras con una resolución muy alta, por lo que se ajusta mucho a los datos pero es posible que sobre-ajuste; *Medium Gaussian* utiliza fronteras con un valor intermedio de resolución; y *Coarse Gaussian* utiliza fronteras más suaves, por lo que es más robusto frente a datos ruidosos pero puede infra-ajustar.

- **Clasificador de vecinos más cercanos (k-NN)**

El clasificador de vecinos más cercanos es un modelo que no aprende durante el entrenamiento, sino que simplemente almacena los datos, se seleccionan k vecinos inicialmente y para clasificar un nuevo punto, el modelo observa las clases de los k vecinos más cercanos y le clasifica en la categoría a la que pertenezcan la mayoría de dichos vecinos [52].

Matlab ofrece 3 modelos en función del número de vecinos: *Fine KNN*, que utiliza pocos vecinos, ajustando mucho la frontera de decisión, por lo que es muy sensible a datos ruidosos; *Medium KNN* que utilizan un número intermedio de vecinos, por lo que equilibra precisión y suavidad de la frontera; y por último *Coarse KNN*, el cual utiliza un gran número de vecinos y por tanto utiliza una frontera de decisión más suave pero menos ajustada.

Por otro lado, también se ofrecen 2 modelos en función de la distancia, estos son *Cosine k-NN*, que mide el ángulo entre dos vectores y les considera cercanos cuando apuntan en

la misma dirección; y *Cubic k-NN*, que calcula la distancia siguiendo un polinomio de tercer grado.

Por último, existe el modelo *Weighted KNN*, que tiene en cuenta todos los puntos, pero pondera en mayor grado los vecinos cercanos a los lejanos [52,53].

- **Clasificadores combinados**

Consisten en ensamblar varios métodos clasificadores, de forma que se obtiene un resultado más exacto a partir de la combinación de los resultados.

Se ofrecen modelos combinados de varios tipos en Matlab, en primer lugar árboles de decisión, los cuales existen *boosted trees*, *bagged trees* y *RUSBoosted trees*.

Boosted trees consiste en un proceso iterativo que va creando árboles secuencialmente, cada nuevo árbol da mucha importancia a los errores cometidos por los árboles anteriores, de esta forma se crean muchos árboles poco precisos, los cuales al ensamblarse forman un modelo muy preciso [54,55].

RUSBoosted trees, antes de entrenar cada árbol, elige aleatoriamente un número de ejemplos similar al de la clase minoritaria, de forma que todas las clases tienen aproximadamente el mismo número de ejemplos. Una vez hecho esto, funciona de la misma forma que *Boosted trees*.

Bagged trees, a diferencia de *boosted trees*, que tenía en cuenta los resultados de cada árbol para entrenar los siguientes, en este caso cada árbol se entrena de forma independiente a los demás utilizando una parte de los datos escogida de forma aleatoria, los resultados se obtienen realizando el promedio de los resultados de todos los árboles [56].

Por otra parte se encuentran los modelos *Subspace Discriminant* y *Subspace KNN*, ambos funcionan de la misma forma pero utilizando métodos distintos. Es decir, ambos eligen aleatoriamente un subconjunto de características para entrenar en cada iteración y se elige la solución escogiendo la predicción que más se ha predicho. Pero se diferencian en que *Subspace Discriminant* utiliza los clasificadores lineales de análisis discriminante explicados anteriormente, mientras que *Subspace KNN*, utiliza los de tipo vecinos más cercanos del apartado anterior.

- **Redes neuronales**

Las redes neuronales, consisten en modelos formados por neuronas que se comunican mediante conexiones ponderadas, y se agrupan en capas, de forma que las neuronas de cada capa toman como entrada características iguales. Estas capas se dividen en 3 grupos: capa de entrada, que recibe los datos iniciales; capas ocultas, puede ser una o varias capas y procesan la información de forma no lineal, lo que hace que las redes neuronales sean tan eficaces identificando patrones; y capa de salida, formada por una única neurona que genera el resultado final [57].

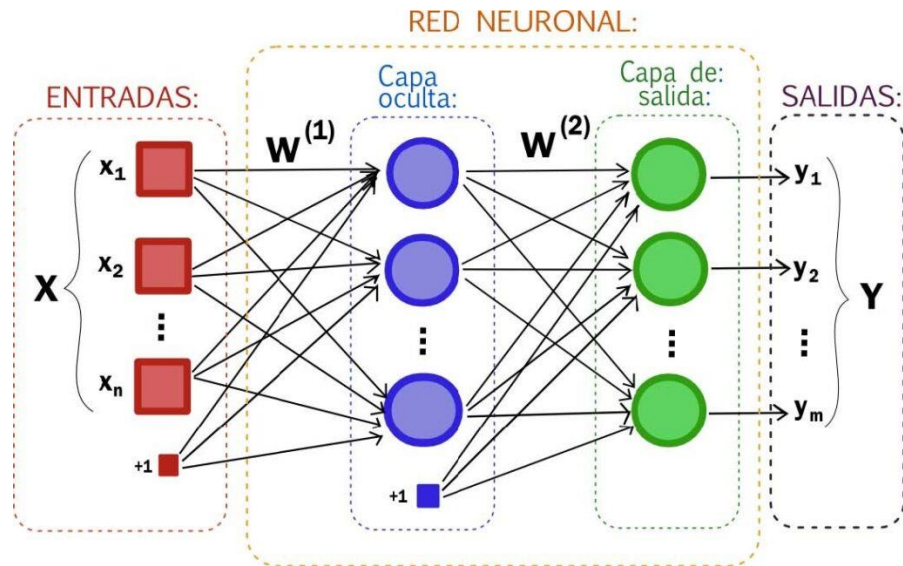


Figura 38. Esquema de red neuronal [58].

En Matlab hay disponibles varios tipos de redes neuronales predefinidas en función del tamaño de estas, en primer lugar, *Bilayered Neural Network* y *Trilayered Neural Network*, las cuales tienen únicamente dos y tres capas respectivamente (sin contar la capa de entrada). Por otro lado se pueden crear las redes *Narrow Neural Network*, *Medium Neural Network* y *Wide Neural Network*, ordenadas de menor a mayor tamaño respectivamente. Por último, se puede elegir *Optimizable Neural Network* y seleccionar a mano los parámetros de la red neuronal como son el tamaño de las capas, el número de capas y la tasa de aprendizaje entre otros.

Cuanto más grande sea la red neuronal que se ha elegido mayor será su precisión pero también será mayor su coste computacional. Es por ello que tiene una gran importancia la elección del tamaño de la red neuronal y con ello su número de capas.

5.3. Elección del modelo utilizado

Una vez vistos todos los modelos posibles, en el siguiente apartado se procederá a su entrenamiento y se comprobará cuáles ofrecen mejores resultados. Se eliminarán algunos de los modelos a medida que se vayan realizando pruebas, para finalmente decidir cuál será el modelo utilizado para el estudio.

Se decidirá qué modelo es el mejor para este trabajo en base a los siguientes criterios: su precisión clasificando espectros en las especies entrenadas, si son capaces de clasificar espectros desconocidos, su efectividad clasificando individuos de la misma especie y su coste computacional.

6. Desarrollo de un programa en Matlab para la distinción de especies usando los espectros obtenidos con LIBS.

6.1. Introducción y objetivo del desarrollo

Como ya se ha comentado anteriormente, la técnica LIBS posee un gran potencial para múltiples aplicaciones. Además, cuando a los datos obtenidos se les aplica algún tipo de procesamiento de señal, dicho potencial se incrementa de forma notable, permitiendo abordar tareas complejas. En este trabajo se explora específicamente la discriminación de especies animales a partir del análisis de muestras huesos mediante LIBS.

Esta tarea puede ser de gran utilidad para analizar restos de huesos encontrados, distinguir huesos que se encuentren en fosas comunes y otras numerosas utilidades, por ejemplo de ámbito forense [34]. Además, se podría mejorar esta técnica en un futuro para distinguir individuos de una misma especie o incluso para utilizarla en labores de autopsias con huesos humanos.

Para este trabajo de fin de grado se utilizarán dos huesos de cordero, uno de cabra y dos de ciervo, a los cuales se les realizarán medidas tanto de forma lineal, las cuales consisten en seguir una línea con el láser, a través de la que se va disparando a la muestra y recogiendo los datos; como usando mapeo, que se diferencia de las medidas lineales en que en vez de una recta se escoge una superficie se realiza una malla en dicha superficie y se dispara el láser en todos los puntos de la malla. Una vez realizadas estas medidas y recopilados los espectros, se entrenará una red neuronal con los datos para que al pasarle un espectro sea capaz de distinguir a cuál de las especies nombradas anteriormente pertenece.

A continuación, en las figuras 39 y 40, se pueden observar dos ejemplos de medidas tanto por mapeo (2 dimensiones) como de forma lineal (1 dimensión).



Figura 39. Ejemplo de media por mapeo. Elaboración propia.



Figura 40. Ejemplo de media lineal. Elaboración propia.

En las figuras 41 y 42 se pueden observar los huesos de cordero (C) y el de cabra (CA), antes y después de ser recogidos, ya que estos huesos han sido enterrados para simular las condiciones que habrían recogido si se hubiesen encontrado enterrados. Además, en la figura 43, se ve el hueso de ciervo rojo recogido en el monte.



Figura 41 . Huesos de cordero y cabra enterrados. Elaboración propia.



Figura 42 . Huesos de cordero y cabra tras ser recogidos y preparados para su análisis en el laboratorio.
Elaboración propia.

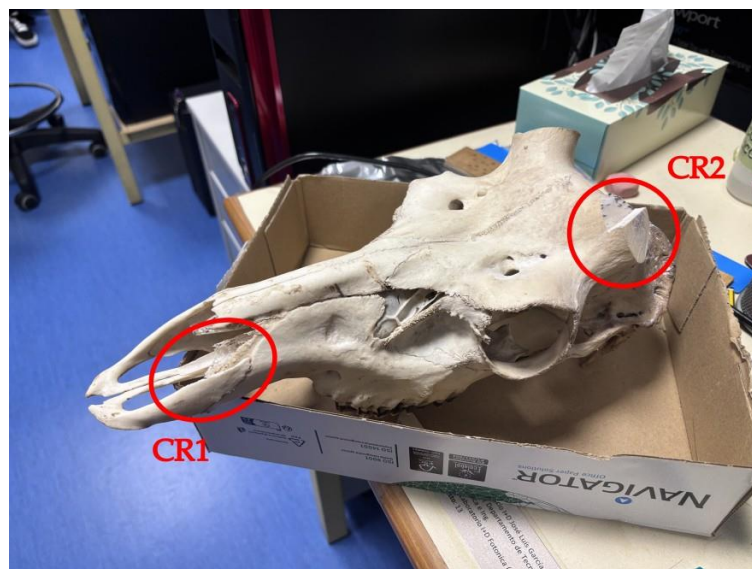


Figura 43 . Hueso de ciervo rojo recogido en el monte. Elaboración propia.

A este cráneo de ciervo rojo, se le han extraído dos muestras (CR1 y CR2), las cuales son las utilizadas para el trabajo, estas se pueden observar en la figura 44.



Figura 44 . Muestras CR1 y CR2 extraídas del cráneo de ciervo rojo. Elaboración propia

Por otra parte, se tienen varias muestras de huesos encontrados en el monte, los cuales no se sabe a qué especies pertenecen, serán de utilidad para comprobar de qué forma les clasifica la red que se implemente en este trabajo. Estas pueden verse en las figuras 45, 46 y 47.



Figura 45 . Muestras desconocidas (D) recogidas en el monte. Elaboración propia.



Figura 46. Muestras desconocidas (Z) recogidas en el monte. Elaboración propia.



Figura 47 . Muestras desconocidas de vertebra (V) recogidas en el monte. Elaboración propia.

En la siguiente tabla se muestran todos los huesos disponibles para el trabajo.

Etiqueta	Especie	Origen
CA	Cabra	Enterrado
C	Cordero	Enterrado
CR	Cráneo Ciervo 1	Monte
CRA	Cráneo Ciervo 2	Monte
D	Desconocido 1	Monte
Z	Desconocido 2	Monte
V	Vertebra (Desconocido 3)	Monte
O	Omoplato (Desconocido 4)	Monte

Tabla 1. Huesos disponibles para el trabajo. Elaboración propia.

6.2. Medidas realizadas

En primer lugar, se han tomado los dos huesos de cordero y se les ha realizado 4 medidas mapeadas y 2 lineales a cada uno. Después, se le han realizado 3 medidas mapeadas y 1 lineal al hueso de cabra, y por último se han realizado 2 medidas por mapeo al primer hueso de ciervo y 3 al segundo. Por lo que en total se han realizado 21 medidas.

En la figura 48, se ve el hueso de cordero 1 (C1), tras realizarle un mapeado con el láser LIBS, se puede ver el efecto provocado por esta técnica, ablacionando una parte de la muestra.



Figura 48. Efecto provocado por LIBS a la muestra C1. Elaboración propia.

Para obtener las medidas se realizan 15 disparos en cada punto, de forma que los primeros disparos sirven para eliminar la primera capa de la muestra, la cual puede estar contaminada por suciedad y por otros elementos con los que habrá estado en contacto el hueso en caso de haber estado enterrada (por ejemplo). Mientras que los siguientes disparos dan lugar a las medidas que se utilizan. Concretamente se han eliminado los 6 primeros disparos en cada punto para todas las muestras y se han mantenido los 9 últimos. En la siguiente imagen se puede observar cómo va cambiando el espectro recogido de un mismo punto a medida que se le va haciendo disparos con el láser.

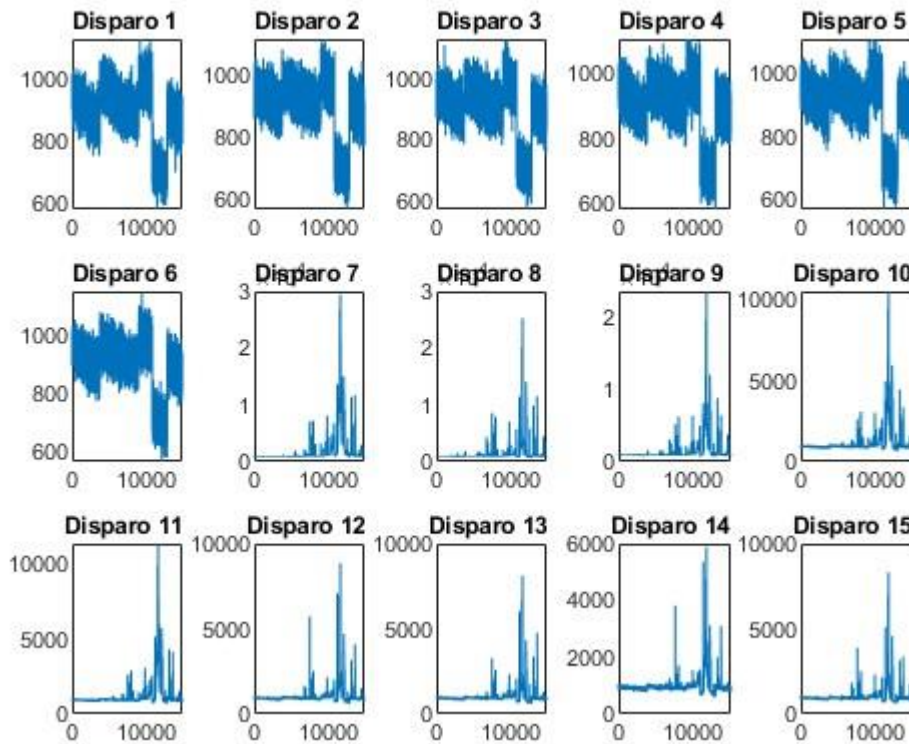


Figura 49. Evolución del espectro en un mismo punto con 15 disparos para el hueso de cordero 1. Elaboración propia.

En la figura 49 se puede observar lo explicado anteriormente, los 6 primeros disparos tienen una forma que no se parece en nada a lo que se busca, esto es debido, como se ha explicado anteriormente, a que las muestras tienen algo de suciedad en la superficie. De esta forma, como la técnica LIBS elimina una pequeña parte del material, en estos 6 primeros disparos lo que se está haciendo es eliminar dicha suciedad, y a partir del disparo 7, cuando ya no queda suciedad en la muestra, se empiezan a recoger los espectros útiles para el trabajo.

Estas medidas se han cargado en Matlab, obteniendo 21 matrices, una para cada medida realizada. Siendo las matrices de las medidas lineales de 3 dimensiones, ya que se tiene un espectro para cada una de las 15 medidas en cada uno de los puntos, y las matrices de las medidas mapeadas de 4 dimensiones, ya que además de las 3 dimensiones anteriores hay que añadir una más por realizarse las medidas en una superficie y no en una recta. Para poder tratar los datos de forma más sencilla, se han eliminado los 6 primeros disparos en cada punto como se ha dicho anteriormente, y se han re-escalado todas estas matrices, convirtiéndolas en matrices de dos dimensiones, de forma que cada fila sea un espectro y cada columna contenga las longitudes de onda, sin tener en cuenta de que punto del hueso se ha obtenido cada espectro.

Por otro lado, se han normalizado las medidas entre 0 y 1, ya que puede haber diferencias en la configuración utilizada para el láser entre unas medidas y otras, lo que podría hacer que los datos no fueran fiables.

Una vez ya se tienen todos los datos normalizados y en matrices de dos dimensiones, se han juntado las matrices de cada muestra, dando lugar a 5 matrices distintas. Como los datos obtenidos son demasiados, si se pasasen esas matrices directamente a una red

neuronal, los tiempos de entrenamiento serían demasiado elevados, además la red se sobre-ajustaría y no funcionaría correctamente. Para solucionar dicha problemática se han eliminado una parte de los datos aleatoriamente. Quedando finalmente las siguientes 5 matrices:

- **EspectroC1L1_2Dnorm**, contiene los espectros del hueso de cordero 1, tiene dimensiones 3690x15049.
- **EspectroC2L1_2Dnorm**, son los espectros del hueso de cordero 2, con dimensiones 1665x15049.
- **EspectroCAL1M1_2Dnorm**, contiene los espectros obtenidos del hueso de cabra, tiene dimensiones 10738x15049.
- **EspectroCR1M1_2Dnorm**, están guardados los espectros del hueso de ciervo 1, con dimensiones 5000x15049.
- **EspectroCR2M1_2Dnorm**, guarda los espectros del hueso de ciervo 2, y la cual tiene dimensiones 5000x15049.

6.3. Desarrollo del programa

6.3.1. Almacenamiento de las medidas y representación de los espectros

El programa que se ha creado, carga en primer lugar los datos medidos y los guarda en las matrices explicadas en el apartado anterior. Además, para cada una de las muestras, se hace la media de todos los espectros medidos y se dibuja el espectro medio calculado, de forma que el usuario del programa pueda ver a simple vista las diferencias en los espectros para cada especie antes de entrenar la red neuronal con los datos. Esto también podría ser útil en caso de que las diferencias de espectros entre las especies que se pretendan comparar sean diferenciables a simple vista, y el usuario no quiera entrenar una red neuronal, de modo que podría realizar la distinción de especies a simple vista observando los espectros.

A continuación, en las figuras 50 a 54, se pueden ver los espectros medios obtenidos para las 5 muestras de huesos que se han medido. En primer lugar, en las figuras 50 y 51, se pueden observar dos espectros de las dos muestras de huesos de cordero (figura 42) que se han medido. Se puede ver que salvo alguna diferencia en las intensidades en algunas longitudes de onda, sus formas son muy similares.

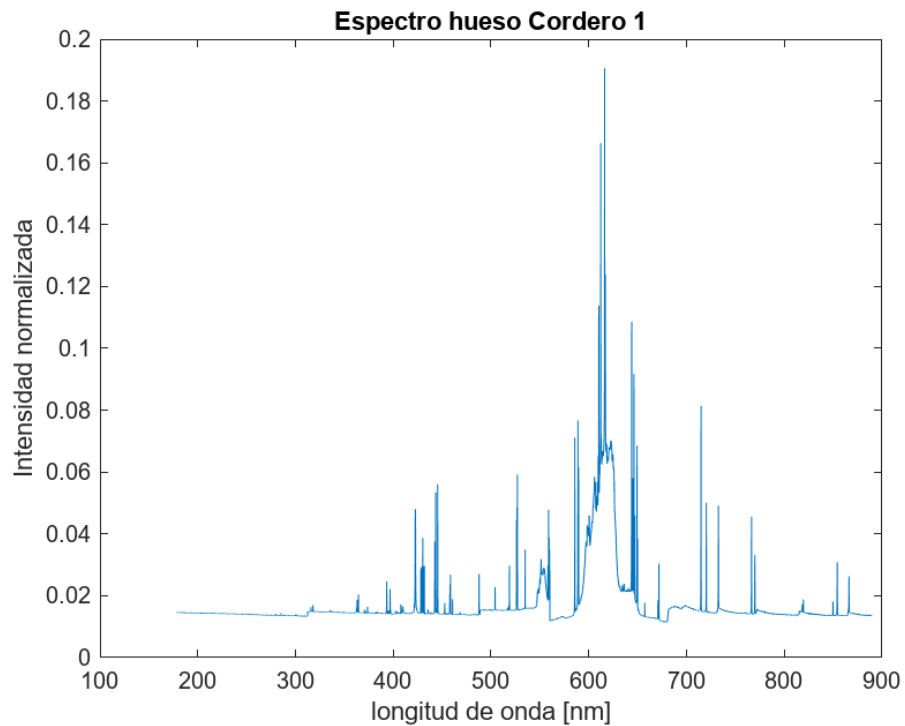


Figura 50. Espectro obtenido con LIBS en el hueso de cordero 1. Elaboración propia.

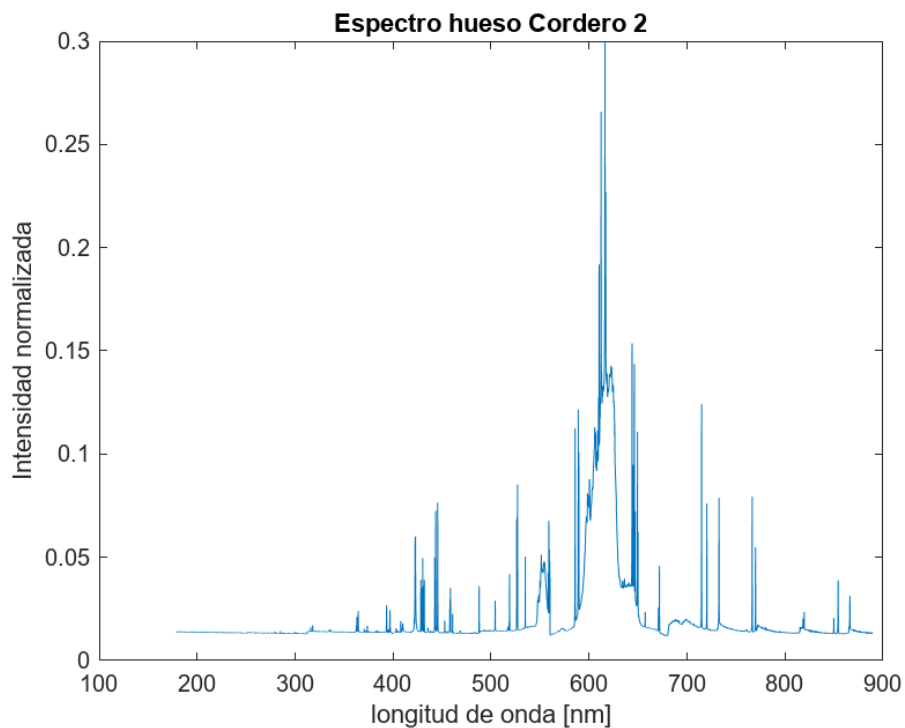


Figura 51. Espectro obtenido con LIBS en el hueso de cordero 2. Elaboración propia.

Seguidamente, en la figura 52 se puede observar el espectro obtenido para el hueso de cabra, en este caso solo se tiene una muestra para esta especie, por lo que solo se ve un espectro. No obstante, se puede ver que se diferencia claramente de los espectros de cordero, pudiendo diferenciarlos incluso a simple vista.

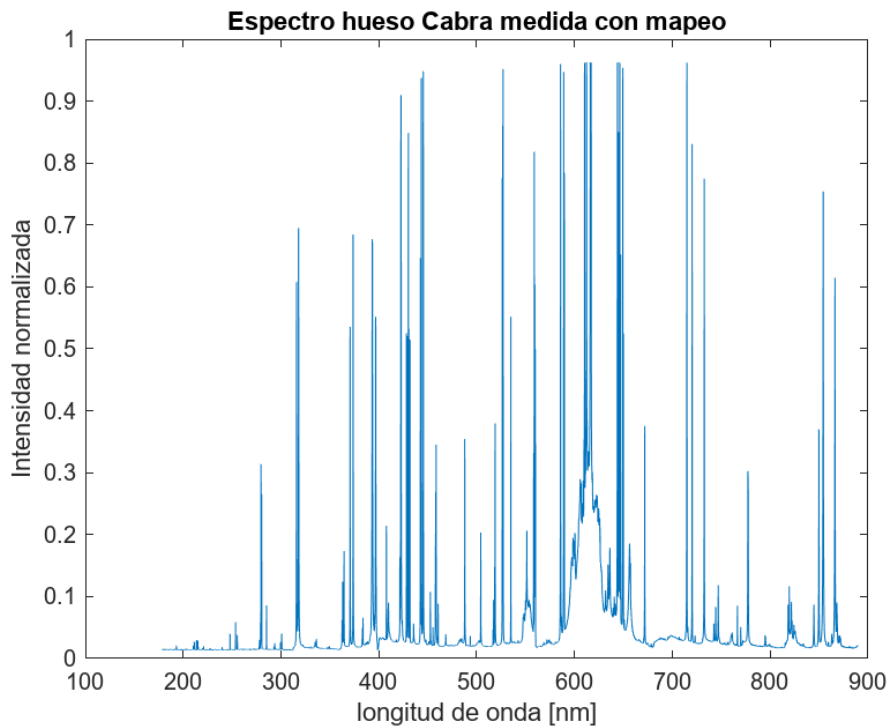


Figura 52. Espectro obtenido con LIBS en el hueso de cabra. Elaboración propia.

Por último, en las figuras 53 y 54 se tienen los dos espectros obtenidos para las dos muestras de ciervo medidas. En este caso aunque sigue siendo bastante distinto del espectro de cordero, no se diferencia en tanta medida con el espectro de cabra. Por otro lado, como ocurría anteriormente, aunque sean distintas muestras, ambos espectros son muy similares.

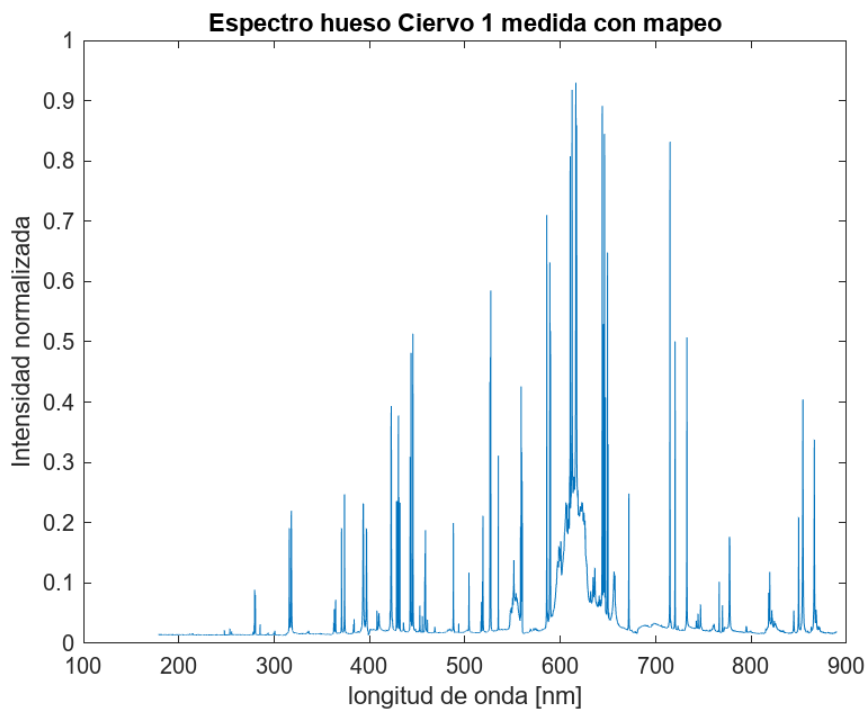


Figura 53. Espectro obtenido con LIBS en el hueso de ciervo 1. Elaboración propia.

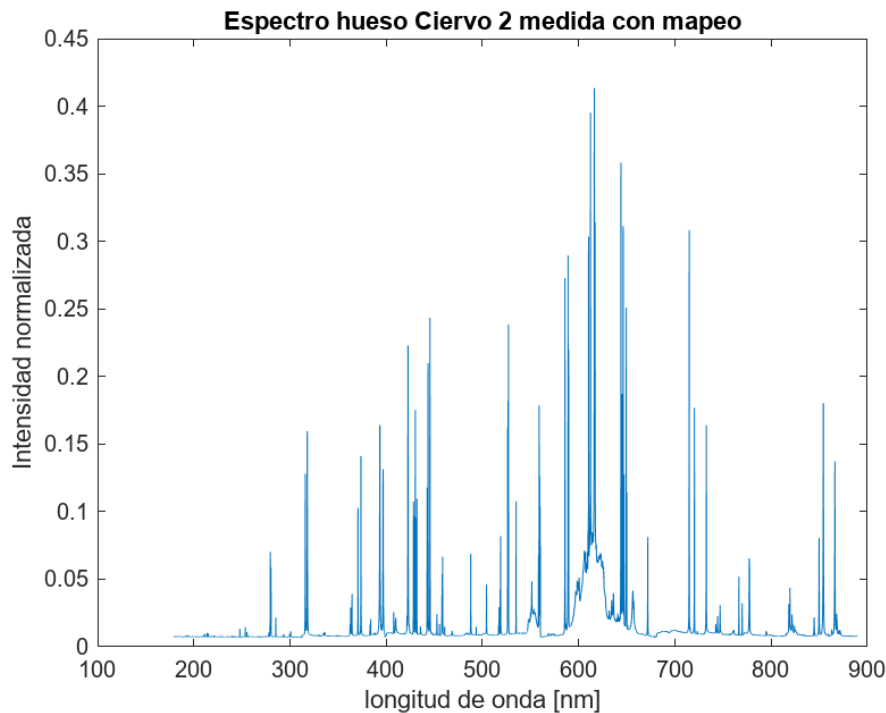


Figura 54. Espectro obtenido con LIBS en el hueso de ciervo 2. Elaboración propia.

A continuación se muestra como ejemplo el código utilizado para cargar la matriz y dibujar el espectro para la muestra C1L1, el resto del código puede verse en el anexo 1.

Muestra C1L1 (L = lineal, C1 = Cordero 1)

```
cd 02062025-133541_MedidaHueso_C1_L1_pumpenergy18
load ('matlabData.mat','spectra')
Espectro = spectra;
EspectroC1L1 = Espectro(:,7:15,:);
%Se eliminan las primeras 6 medidas en cada punto porque sirven para
quitar
%suciedad de la muestra y no proporcionan datos reales
EspectroC1L1_2D = reshape(EspectroC1L1, size(EspectroC1L1,1), []);
EspectroC1L1_2D = EspectroC1L1_2D';

%Se han normalizado los datos para poder utilizar datos de diferentes
especies y medidas con diferentes potencias del laser
EspectroC1L1_2Dnorm = rescale(EspectroC1L1_2D, 0, 1); %Normaliza entre 0
y 1
tamano = size(EspectroC1L1_2D);
mediaC1L1 = (sum(EspectroC1L1_2Dnorm,1))/(tamano(1));
figure;plot(mediaC1L1); title('Espectro hueso Cordero 1')
```

6.3.2. Separación de los datos en entrenamiento y validación y creación de las matrices asociadas

En este punto, los datos que se tienen para entrenar la red neuronal se dividen en dos grupos, unos para entrenar la red y otros para comparar las predicciones de las redes generadas con otros datos que la red neuronal no haya visto antes, ya que si no podría

ocurrir que la red predijese bien los resultados con los datos con los que ha sido entrenada, pero no funcionase correctamente cuando se aportasen nuevos datos. Concretamente se van a utilizar el 80% de los datos que se tienen para el entrenamiento de la red, y el otro 20% para la validación de los resultados.

Para ello, se utilizan 5 matrices de entrenamiento, llamadas *EntrenamientoClaseX*, numeradas de la 1 a la 5, siendo X el número de matriz, en cada una de las cuales se guardan los datos de una muestra distinta. Tras esto, se hace la separación de los datos explicada anteriormente, así que se guardan el 80% de los datos en cada matriz de entrenamiento, y el otro 20% en matrices llamadas *TestClaseX*, que servirán para validar los resultados cuando se haya creado la red neuronal.

De esta forma, se unen las 5 matrices denominadas *EntrenamientoClaseX* en una única matriz, dando lugar a la matriz, *MatrizEntrenamiento*, en la que están contenidos todos los datos que se van a utilizar para entrenar el modelo. Pero para entrenar la red neuronal, además de la matriz con los datos de las medidas, es necesario disponer de una matriz de validación, que clasifique las muestras en las categorías que la red debe diferenciar. En este caso, se establecen tres categorías:

Categoría 1: clases 1 y 2, correspondientes a dos muestras distintas de cordero.

Categoría 2: clase 3, correspondiente al hueso de cabra.

Categoría 3: clases 4 y 5, correspondientes a dos muestras de huesos de ciervo.

Esta clasificación se representará en una matriz denominada Y, de forma que cada fila de *MatrizEntrenamiento* tenga asociada una fila en Y. En dicha fila de Y aparecerá el valor 1, 2 o 3, en función de la categoría a la que pertenezca la muestra correspondiente en la matriz de entrenamiento.

6.3.3. Diseño e implementación de la red neuronal

Una vez se tienen los datos guardados y se han dibujado los espectros, se pasa a generar una red neuronal que sea capaz de teniendo un espectro generado por LIBS, distinguir si pertenece a un hueso de cordero, cabra o ciervo. Para generar la red neuronal, se va a utilizar la aplicación “*classification learner*” de Matlab, ya que ofrece diversos tipos de redes neuronales que podrán ser probados con los datos que se tienen, escogiendo la red que mejor se ajuste a los resultados buscados.

La aplicación *classification learner* de Matlab ofrece crear modelos de los siguientes tipos: análisis discriminante, clasificador Bayesiano, clasificador por regresión logística, máquinas de vectores de soporte (SVM), clasificador lineal, clasificadores de vecinos más cercanos (KNN), clasificador por aproximación de kernel, redes neuronales y árboles de decisión. Se puede seguir el proceso utilizado para entrenar un modelo de estos tipos en las siguientes imágenes.

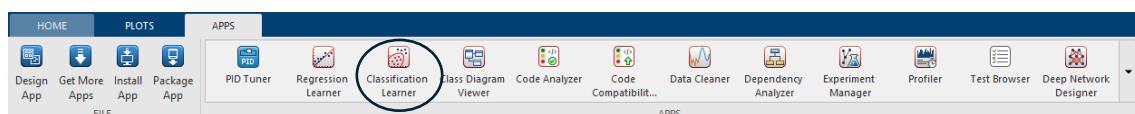


Figura 55. Pestaña de apps de Matlab. Elaboración propia.

En primer lugar se selecciona *Classification Learner* dentro de la pestaña de apps de Matlab. Tras ello, aparece la pestaña que se puede observar en la figura 56, en la cual se debe seleccionar *New Session*.

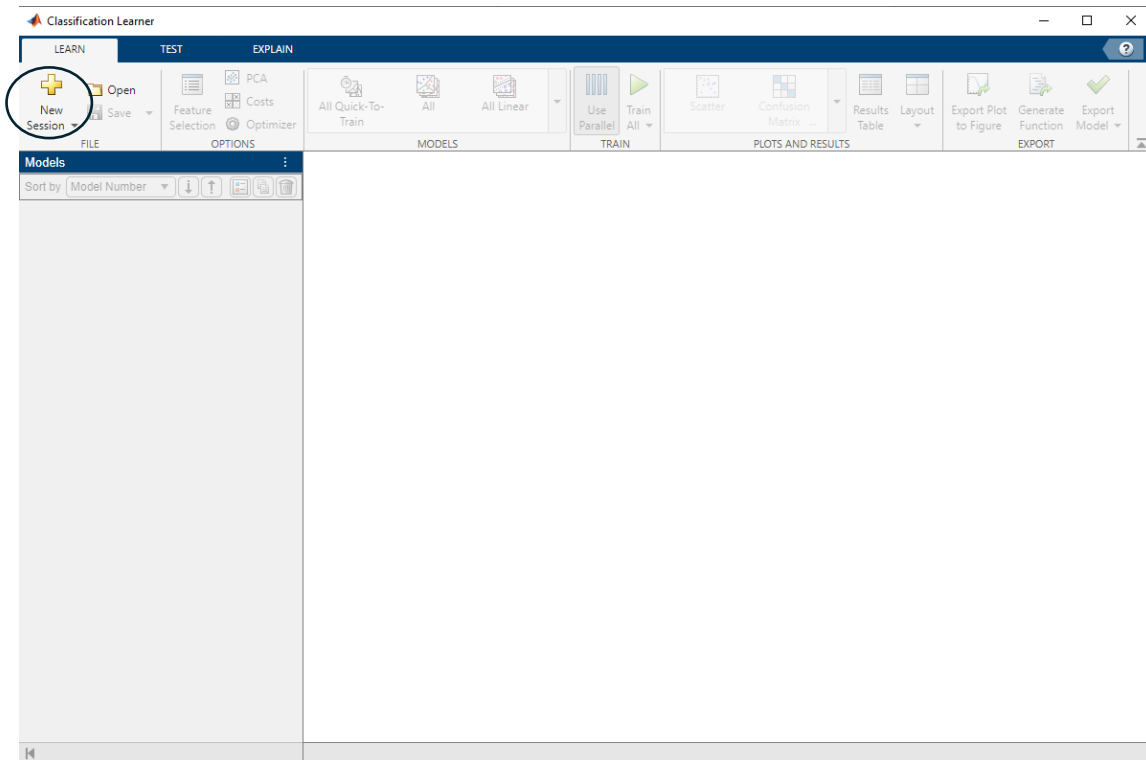


Figura 56. Pestaña de *Classification Learner*. Elaboración propia.

Una vez hecho esto, se llega a la pestaña indicada en la figura 57, en la cual el programa pide al usuario los datos con los que quiere realizar el entrenamiento. En este caso se selecciona *MatrizEntrenamiento*, ya que es la matriz donde están contenidos dichos datos. Además, el programa da la opción de que los datos de validación estén contenidos en una de las columnas de la matriz de los datos de entrenamiento. En este caso, los datos de validación están guardados en la matriz *Y*, por lo que se selecciona la opción *From workspace* dentro de *Response* y se selecciona la matriz *Ycat*, ya que la matriz de validación debe ser de tipo *categorical* por ser modelos de clasificación y no de regresión, la única diferencia de *Y* con *Ycat* es la forma en que se guardan los datos, *Y* es de tipo *double*, mientras que *Ycat* es de tipo *categorical*, pero los valores contenidos son idénticos en ambas matrices.

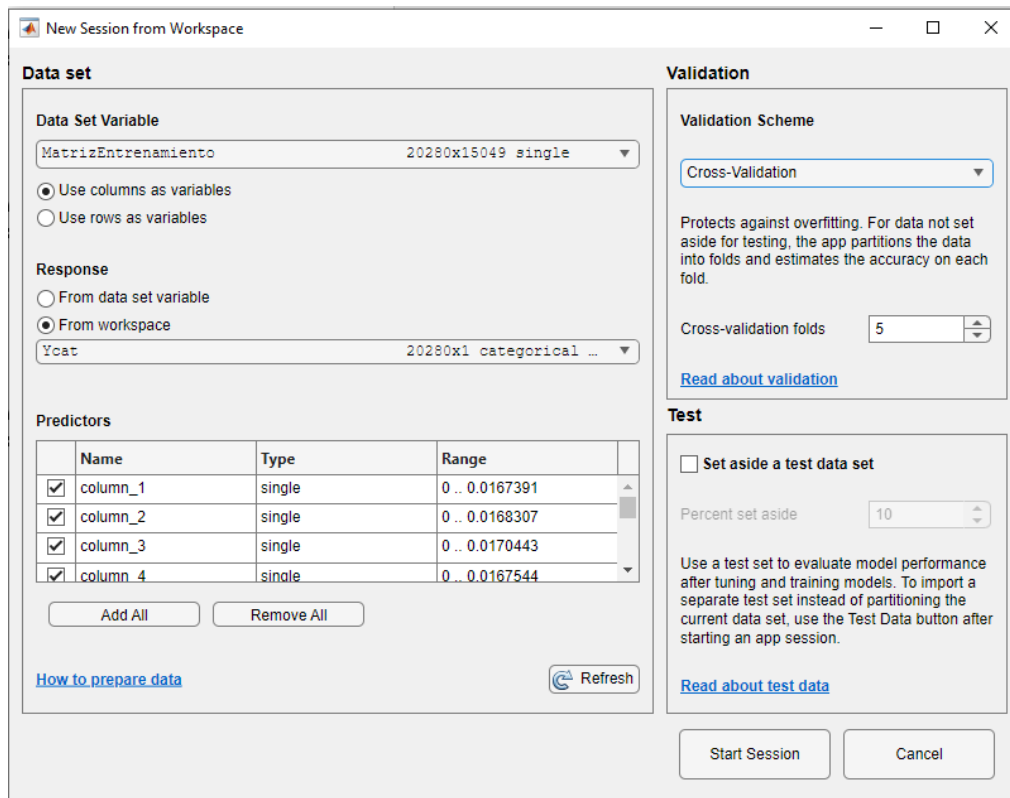


Figura 57. Pestaña de selección de datos para el entrenamiento. Elaboración propia.

Esta interfaz permite además ajustar dos opciones más:

Validation

En la cual se pregunta al usuario la forma en la que quiere que se realice el valor de validación que calcula Matlab al terminar el entrenamiento con cada modelo, en ella se dan 3 opciones.

- *Cross-Validation,*

Divide los datos en k partes que indica el usuario, el modelo se entrena con $k-1$ partes y se realiza la validación con la parte guardada, tras ello, se realiza otra iteración cambiando la parte de validación y se repite el proceso hasta completar las k iteraciones. El valor de validación que indica Matlab después del entrenamiento se calcula realizando la media de los valores de cada iteración. Este es el tipo de validación que se va a utilizar en este trabajo.

- *Holdout Validation*

Se dividen los datos en dos subconjuntos, entrenamiento y validación, se realiza el entrenamiento y se calcula la validación con dicho subconjunto. Es más rápido pero la estimación es peor que con *Cross-Validation*, por lo que es mejor para matrices de datos muy grandes. Además, en este caso ya se ha guardado una parte de los datos para validación, por lo que es mejor utilizar *Cross-Validation* y después realizar la validación con esos datos.

- *Resubstitution Validation*

Entrena y valida con los mismos datos, por lo que es muy común que utilizando esta opción el modelo realice sobre-ajuste, así que también se descarta en este caso.

Test

Sirve para guardar una parte de los datos que no se utilice para el entrenamiento y que solo se use para calcular la validación una vez se haya terminado el entrenamiento. Se podría haber utilizado para este programa, pero esta opción guarda un porcentaje de los datos sin tener en cuenta cuales está guardando. Mientras que en el programa se quiere guardar el mismo porcentaje de datos de cada muestra para verificar que el modelo funciona de forma correcta con todas las muestras, es por ello que los datos de *test* ya han sido guardados anteriormente en las matrices *TestClaseX*.

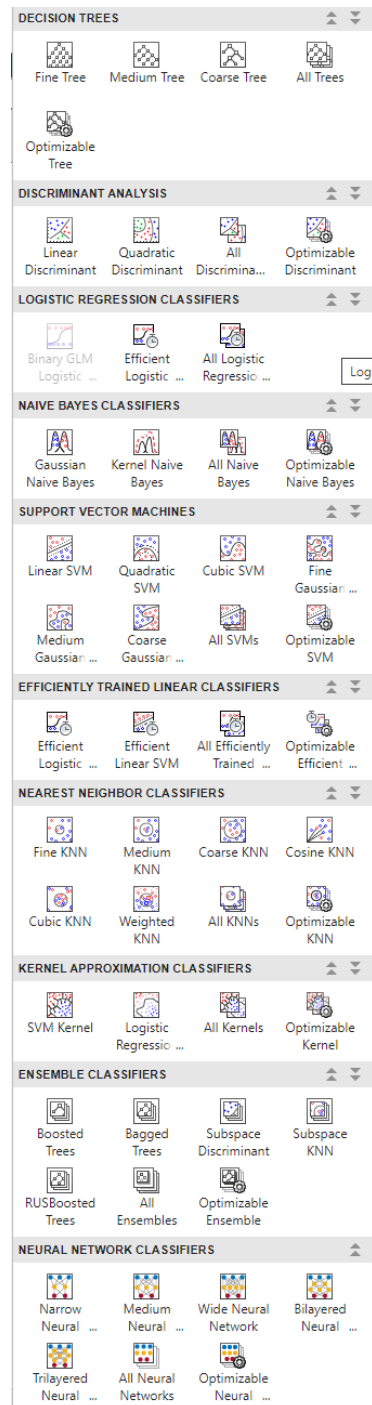


Figura 58. Modelos seleccionables para el entrenamiento. Elaboración propia.

Una vez seleccionadas todas las opciones, se selecciona *Start Session*, llegando a la pestaña de la figura 58, en la que se muestran todos los modelos que se pueden entrenar. Como se puede observar, en dicha pestaña aparecen todos los modelos que se han nombrado al inicio del apartado 5.3.3. Se seleccionan todos los modelos que se quieran entrenar y se pulsa la opción *Train all* que aparece en la figura 56. De este modo se entrenarán dichos modelos y a la finalización del entrenamiento, Matlab indicará el valor de validación, como se puede comprobar en la figura 59.

☆ 2.1 Tree	Accuracy (Validation): 98.8%
Last change: Fine Tree	15049/15049 features
☆ 2.2 Tree	Canceled
Last change: Medium Tree	15049/15049 features
☆ 2.3 Tree	Canceled
Last change: Coarse Tree	15049/15049 features
☆ 2.4 KNN	Accuracy (Validation): 82.7%
Last change: Fine KNN	15049/15049 features
☆ 2.5 KNN	Canceled
Last change: Medium KNN	15049/15049 features
☆ 2.6 KNN	Canceled
Last change: Coarse KNN	15049/15049 features
☆ 2.7 KNN	Canceled
Last change: Cosine KNN	15049/15049 features
☆ 2.8 KNN	Canceled
Last change: Cubic KNN	15049/15049 features
☆ 2.9 KNN	Canceled
Last change: Weighted KNN	15049/15049 features
☆ 2.10 Efficient Logistic Regression	Accuracy (Validation): 96.0%
Last change: Efficient Logistic Regression	15049/15049 features
☆ 2.11 Efficient Linear SVM	Accuracy (Validation): 98.8%
Last change: Efficient Linear SVM	15049/15049 features

Figura 59. Resultados del entrenamiento. Elaboración propia.

Se ha probado a realizar la tarea con todos los tipos de modelos, y en una primera selección se ha descartado el método de clasificador lineal por ser demasiado sencillo y no aportar resultados fiables y los métodos, *análisis discriminante*, *clasificador Bayesiano* y *clasificador de kernel* por tener demasiado coste computacional para realizar la tarea deseada.

Se ha realizado el entrenamiento utilizando los modelos KNN, regresión logística, SVM, árboles de decisión y red neuronal. Dentro de estos, se seleccionaron *Fine Tree*, *Fine KNN*, *Efficient Logistic Regression* y *Efficient Linear SVM*, por presentar la mejor relación entre coste computacional y resultados, mientras que los parámetros de la red neuronal se ajustaron manualmente.

Tras el entrenamiento con estos cinco modelos, MATLAB proporciona la precisión obtenida para cada uno: 98,8% para el árbol de decisión, 82,7% para KNN, 96% para regresión

logística, 98,8% para SVM y 96,2% para la red neuronal. Cabe destacar que estas precisiones se calcularon sobre los datos de entrenamiento, por lo que para obtener valores más fiables se utilizarán los datos reservados para validación. En función de estas pruebas, se descarta el modelo KNN por su relativamente baja precisión. Posteriormente, al evaluar los modelos con los datos de test, se descarta también la regresión logística debido a que, aunque tiene un error del 3,51%, el cual es un error bastante bajo, se van a utilizar otros modelos que tienen mayor porcentaje de acierto. Por su parte, el árbol de decisión y el modelo SVM presentan errores de 1,42% y 1,83%, respectivamente, por lo que se mantienen como modelos finales.

Como ejemplo, se muestra la matriz de confusión obtenida en el entrenamiento del árbol de decisión, se puede observar en la figura 60.

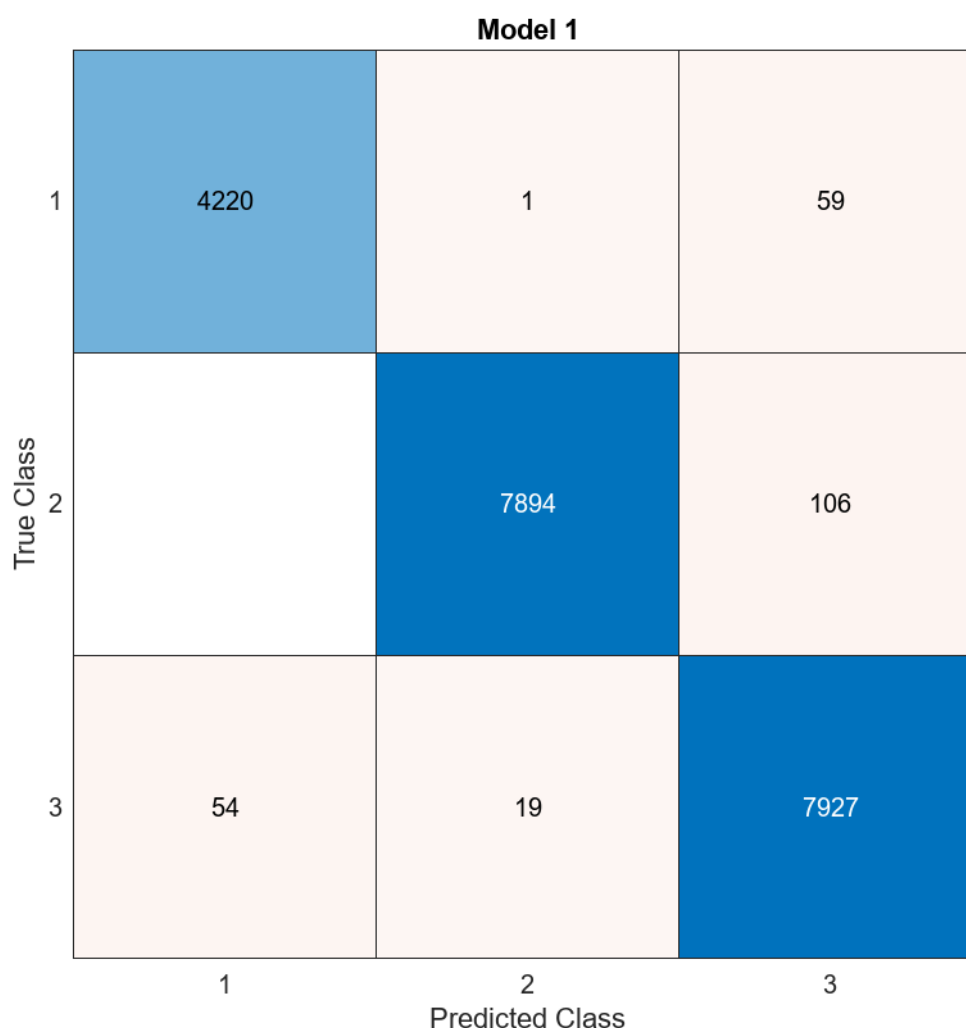
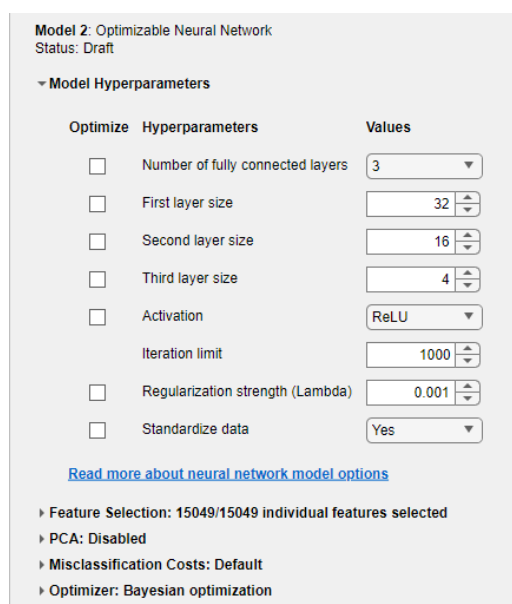


Figura 60. Matriz de confusión en el entrenamiento del árbol de decisión. Elaboración propia.

Por otro lado, el error de la red neuronal es del 3,21%; sin embargo, al poder ajustar sus parámetros manualmente, se mantiene este modelo con el objetivo de intentar mejorar su precisión o de emplearla en el futuro para la clasificación de otro tipo de muestras o en una aplicación similar.

Se ha intentado crear una red neuronal que ofrezca mejores resultados mediante el ajuste de sus parámetros en la interfaz de Matlab, se puede seguir el proceso en la figura 61, en la cual se puede ver que tras varios intentos de entrenamiento de la red neuronal, los mejores

resultados se han obtenido al seleccionar un red con 3 capas ocultas con 32, 16 y 4 neuronas respectivamente, además de la capa de salida, y utilizando una tasa de aprendizaje de 0,001, obteniendo así un error del 2,1% según Matlab, pasando a ser de un 1,9% al comprobarlo con los datos de test.



Model 2: Optimizable Neural Network
Status: Draft

▼ Model Hyperparameters

Optimize	Hyperparameters	Values
☐	Number of fully connected layers	3
☐	First layer size	32
☐	Second layer size	16
☐	Third layer size	4
☐	Activation	ReLU
☐	Iteration limit	1000
☐	Regularization strength (Lambda)	0.001
☐	Standardize data	Yes

[Read more about neural network model options](#)

- Feature Selection: 15049/15049 individual features selected
- PCA: Disabled
- Misclassification Costs: Default
- Optimizer: Bayesian optimization

Figura 61. Parámetros establecidos para la red neuronal. Elaboración propia.

Sin embargo, dado que el coste computacional de esta red es mucho mayor que los modelos que se tenían anteriormente y que sus resultados son incluso algo peores, se ha decidido mantener únicamente los modelos de árbol de decisión y SVM.

6.3.4. Diferenciación entre individuos de la misma especie

Para decidir cuál de estos modelos es mejor, además de la prueba que ya se ha hecho anteriormente para comprobar su acierto diferenciando entre especies, ahora se va a comprobar su acierto distinguiendo entre diferentes individuos de una misma especie. Para ello se han utilizado los datos para las muestras de cordero como se puede observar en el siguiente código.

Una vez ejecutado entrenados los modelos para distinguir entre individuos, Matlab indica los siguientes errores y matrices de confusión durante el entrenamiento.

★ 1 Tree	Accuracy (Validation): 94.1%
Last change: Fine Tree	15049/15049 features
★ 4 Efficient Linear SVM	Accuracy (Validation): 98.7%
Last change: Efficient Linear SVM	15049/15049 features

Figura 62. Precisiones de los modelos SVM y árbol de decisión distinguiendo entre individuos. Elaboración propia.

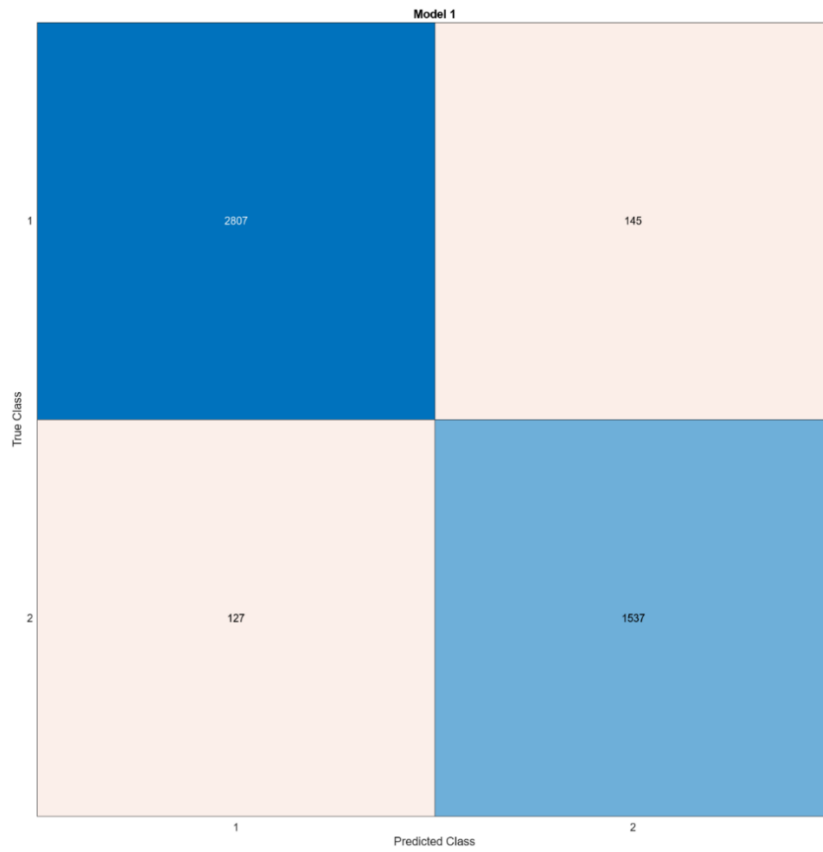


Figura 63. Matriz de confusión durante el entrenamiento con el modelo de árbol de decisión. Elaboración propia.

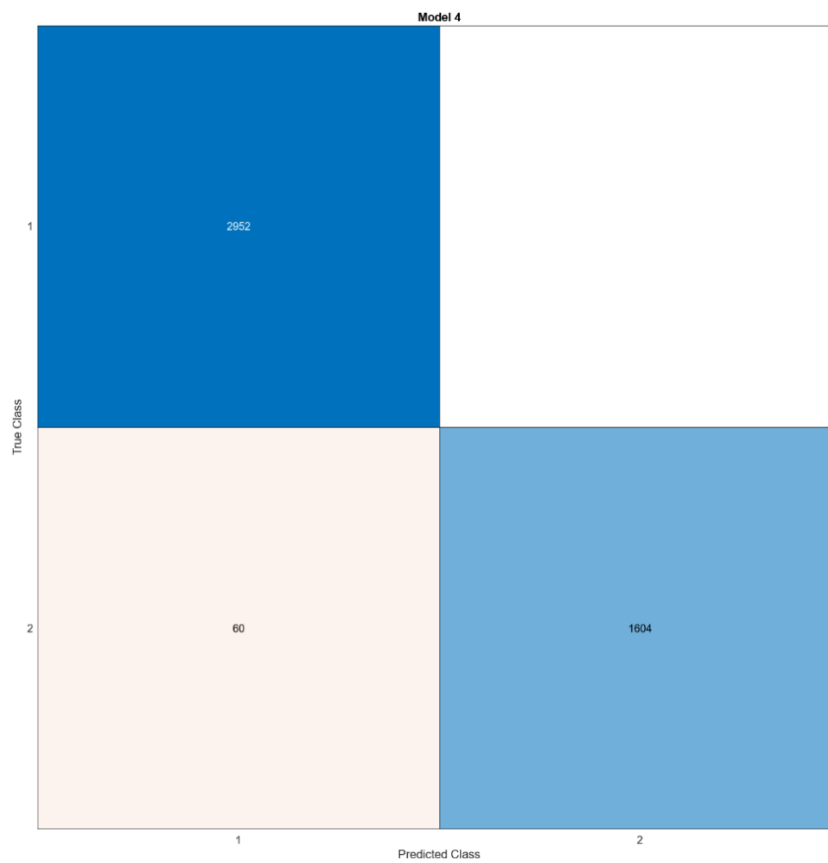


Figura 64. Matriz de confusión durante el entrenamiento con el modelo de SVM. Elaboración propia.

Tras ejecutar el código anterior y medir el error con los datos de test externos al entrenamiento, se obtienen errores de 3,0812% y 1,8674% para el árbol de decisión y SVM respectivamente. Por lo que se concluye que el modelo de árbol de decisión es algo mejor en cuanto a diferenciación de especies (un 0,4% mejor), mientras que el modelo de SVM es un 1,2% mejor al diferenciar entre individuos de la misma especie.

6.3.5. Clasificación de los huesos desconocidos

Ya se ha creado el modelo para realizar predicciones de especies, por lo que se va a probar a pasarle los espectros de las muestras encontradas en el monte para las que no se conoce su especie, observando de qué forma las clasifica la red. Para realizar esto, de entre todos los espectros que se tienen para cada hueso desconocido, se han seleccionado 20000 aleatoriamente, se pasarán estos espectros al árbol de decisión y se contarán cuántos espectros clasifica en cada categoría.

En primer lugar, se cargan los datos del hueso de vertebra (V1) cuya especie se desconoce, para ello se ha utilizado el mismo código que se utilizaba para cargar los datos anteriormente, modificando las variables, este código puede comprobarse en el anexo 1. Una vez cargados los datos, se obtiene el espectro de la figura 65, correspondiente al hueso V1.

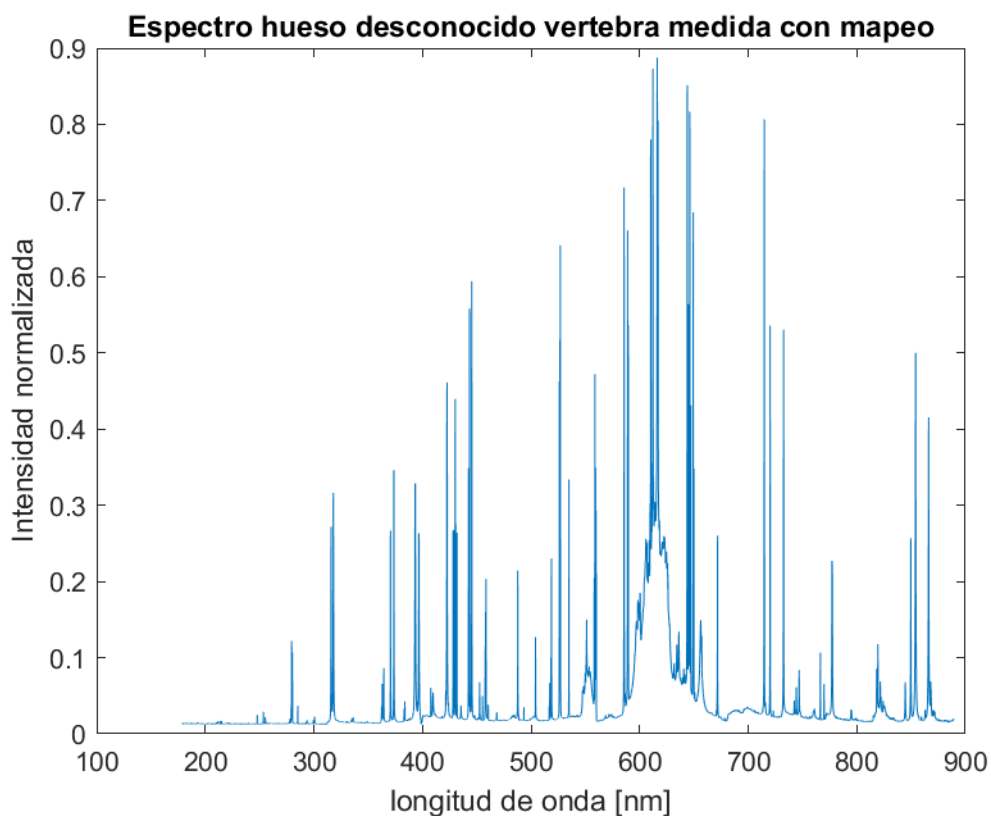


Figura 65. Espectro obtenido para el hueso desconocido V1. Elaboración propia.

Observando el espectro se puede comparar a simple vista con los que se obtuvieron para cada categoría anteriormente en el entrenamiento del árbol de decisión, en esta comparación, parece que el espectro más similar es el de la categoría 3 (ciervo rojo). Tras ejecutar el código y contar cuántos espectros de los 20000 seleccionados clasifica en cada categoría, se obtienen 19 para la categoría 1 (cordero), 3128 para la categoría 2 (cabra) y

16853 para la categoría 3 (ciervo rojo), por lo que se clasifica la muestra como una vértebra de un ciervo rojo.

En segundo lugar, se va a analizar el hueso desconocido 1 (Z1), como en el anterior caso, se cargan los datos y se guardan 20000 espectros seleccionados, descartando el resto para esta prueba, se ejecuta el código con estos datos y se obtiene el espectro de la figura 66.

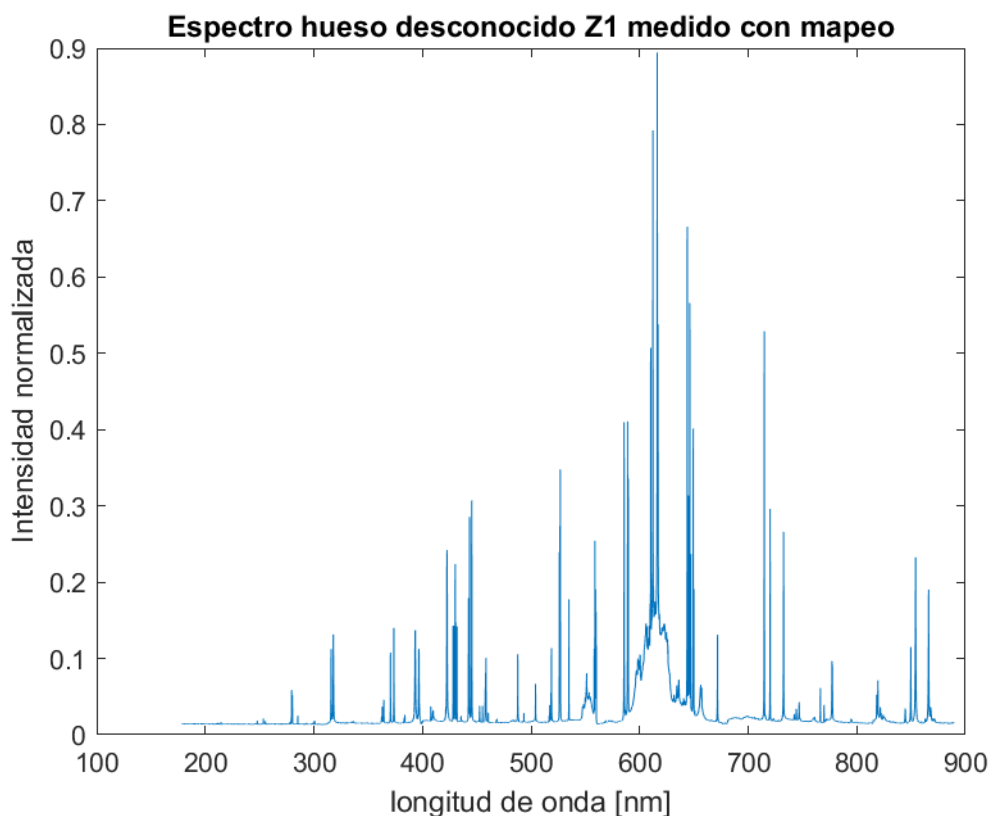


Figura 66. Espectro obtenido para el hueso desconocido Z1. Elaboración propia.

Para este nuevo espectro, a simple vista, aunque no parece tan claro como en el ejemplo anterior, también parece que el espectro más similar es el de ciervo rojo, se comprobará a continuación utilizando el árbol de decisión.

De nuevo se ejecuta el código anterior, y se contabilizan el número de veces que se clasifican los espectros en cada categoría, en este caso, se clasifica 0 espectros como categoría 1, 3949 veces como categoría 2 y 16051 como categoría 3. Por lo que como parecía a simple vista, el hueso se clasifica también como hueso de ciervo rojo.

Por último, se cargan los datos del último de los huesos desconocidos que se posee, el hueso de omoplato (O1), tras seleccionar los 20000 espectros aleatoriamente y ejecutar el código, se obtiene el siguiente espectro, mostrado en la figura 67.

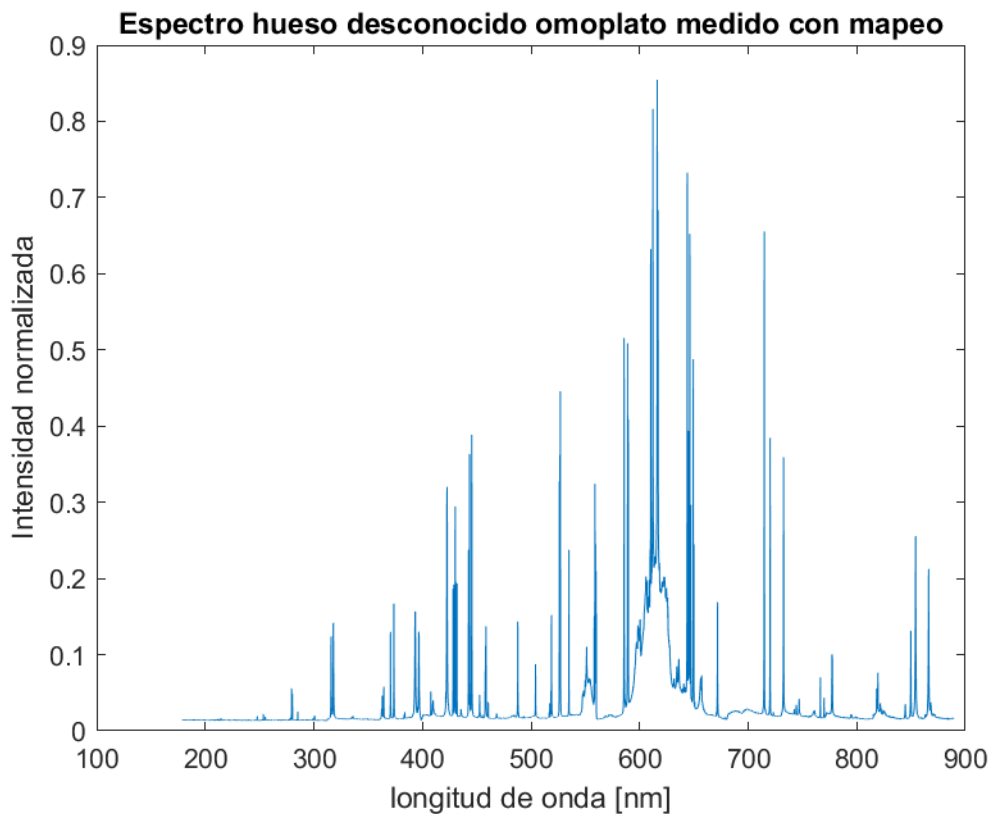


Figura 67. Espectro obtenido para el hueso desconocido O1. Elaboración propia.

A la vista del espectro, este parece prácticamente idéntico al anterior, por lo que a priori parece que el modelo le clasificará de la misma manera.

Una vez pasados los espectros al modelo, se obtienen los siguientes resultados: 378 espectros de categoría 1, 3590 en la categoría 2 y 16032 en la categoría 3. Por ello, esta muestra también queda clasificada como hueso de ciervo rojo.

6.3.6. Conclusión y selección de modelo

Como se ha mostrado a lo largo de este TFG, aunque la técnica LIBS por sí sola ya constituye una herramienta muy útil para este tipo de tareas, su combinación con redes neuronales u otros modelos capaces de interpretar los resultados obtenidos y realizar predicciones con datos futuros la convierte en una herramienta aún más potente.

En relación con los modelos probados en este trabajo, se ha comprobado que existen numerosas alternativas que ofrecen un rendimiento adecuado, si bien cada modelo resulta más eficaz en tareas específicas y algunos presentan un coste computacional demasiado elevado. Por ello, la elección del modelo adquiere una gran importancia y debe realizarse en función de la tarea a abordar.

De los dos modelos que se mantuvieron hasta el final del trabajo, se observa que los resultados obtenidos son bastante similares. No obstante, la selección entre uno u otro depende de si la diferenciación se centra en distinguir especies o en diferenciar individuos dentro de la misma especie. En este caso, se opta por el modelo de árbol de decisión, dado que el objetivo del trabajo es la diferenciación de especies, mientras que la distinción entre individuos se deja como una línea abierta para futuros estudios.

7. Conclusión y líneas futuras

Conclusión

Una vez se ha completado el trabajo, podemos concluir que estamos ante un estudio en el que se han cumplido satisfactoriamente la mayoría de los objetivos establecidos. En un principio no se tenía certeza de si sería posible realizar la distinción de especies utilizando únicamente sus espectros de emisión, obtenidos tras disparar a las muestras con un láser LIBS. Sin embargo, después de llevar a cabo las medidas y tras la construcción del modelo, se comprueba que la tasa de acierto para la diferenciación de los espectros de las tres especies incluidas en el estudio ha sido del 98,58%, lo que permite clasificarlas sin apenas fallos. Asimismo, se descartaron varios modelos con una precisión algo menor, aunque todavía elevada, que podrían haber mostrado un mejor desempeño en otras circunstancias.

En cuanto a las medidas, como se ha observado a lo largo del desarrollo, no es necesario disponer de un número excesivo de las mismas para estas tareas. De hecho, en varias ocasiones ha sido preciso omitir algunas, con el fin de evitar que el tiempo de entrenamiento resultara demasiado prolongado. Esto se debe a que los modelos son capaces de identificar los patrones de los espectros con rapidez, debido a que los correspondientes a diferentes especies presentan diferencias notables, en algunos casos incluso perceptibles a simple vista. Por ello, es importante contar con una matriz de entrenamiento de tamaño suficiente para garantizar la precisión del modelo, pero no tan amplia como para que el proceso de entrenamiento sea ineficiente.

Por otra parte, también han sido descartados algunos modelos como las redes neuronales, modelos SVM o ciertos árboles de decisión similares al utilizado para el trabajo, que podrían haberse mejorado aumentando su complejidad. Sin embargo, en este TFG, se ha priorizado el equilibrio entre el rendimiento del modelo y el coste computacional, por lo que se ha considerado aceptable una tasa de error del 1,42% (que puede considerarse muy baja), dado que su coste computacional no es elevado.

Del mismo modo, se han clasificado tres muestras de especies desconocidas. Al comparar a simple vista sus espectros con los de las especies conocidas, y tras obtener los resultados de como los clasifica el modelo, se comprobó que la red está capacitada para identificar nuevas muestras no vistas anteriormente, siempre que pertenezcan a alguna de las tres especies con las que fue entrenada.

Adicionalmente se ha medido el porcentaje de acierto del modelo para la diferenciación entre individuos de la misma especie (en este caso, cordero), consiguiendo que se distinguiesen con una efectividad de 96,92%. Aunque estos resultados se han logrado modelos como SVM, que proporcionaban mayor precisión en este caso concreto, se ha dado prioridad a la distinción entre especies, por ser el objetivo principal de este trabajo.

Líneas futuras

En primer lugar, como se ha señalado anteriormente, una línea de ampliación de este trabajo podría ser el entrenamiento del modelo para que sea capaz de diferenciar entre individuos de una misma especie. Estos podrían clasificarse en función de diferentes características, como la edad de fallecimiento, zona en la que habitó, su dieta, la presencia

de determinadas patologías u otros datos complementarios que se lograran extraer a partir de los espectros. En esta línea, en el presente trabajo ya se han desarrollado algunas aproximaciones, como el intento de diferenciación entre muestras de (supuestamente) diferentes individuos de cordero, asumiendo que se trataba de huesos de diferentes individuos gracias a los robustos resultados obtenidos por el clasificador.

De manera complementaria a lo anterior, se podría extender el estudio hacia la rama de la medicina forense, creando una herramienta capaz de distinguir los huesos de diferentes sujetos. Anteriormente en este trabajo se han nombrado algunos estudios que realizaban tareas similares a estas [3, 34].

Por otro lado, podría ampliarse el número de especies utilizadas en el entrenamiento, con el fin de obtener un modelo capaz de distinguir entre una mayor cantidad de ellas. Asimismo, podría explorarse la implementación de técnicas de inteligencia artificial que permitan al modelo aprender de manera autónoma a clasificar nuevas especies a medida que se le presenten muestras no conocidas previamente.

En último lugar, se propone la utilización de algunos de los enfoques presentados en algunos de los estudios con muestras de conchas de lapas contemporáneas y arqueológicas que se realizan vía LIBS en el Grupo de Ingeniería Fotónica de la UC, de forma que se tratase de, por ejemplo, clasificar las muestras en diferentes especies (*Patella vulgata*, *Patella depressa*, etc.) en función de sus correspondientes espectros.

Anexo 1: Código Matlab utilizado

Muestra C1L1 (L = lineal, C1 = Cordero 1)

```
cd 02062025-133541_MedidaHueso_C1_L1_pumpenergy18
load ('matlabData.mat','spectra')
Espectro = spectra;
EspectroC1L1 = Espectro(:,7:15,:);
%Se eliminan las primeras 6 medidas en cada punto porque sirven para
quitar
%suciedad de la muestra y no proporcionan datos reales
EspectroC1L1_2D = reshape(EspectroC1L1, size(EspectroC1L1,1), []);
EspectroC1L1_2D = EspectroC1L1_2D';
%Se han normalizado los datos para poder utilizar datos de diferentes
especies y medidas con diferentes potencias del laser
EspectroC1L1_2Dnorm = rescale(EspectroC1L1_2D, 0, 1); %Normaliza entre 0
y 1
tamano = size(EspectroC1L1_2D);
mediaC1L1 = (sum(EspectroC1L1_2Dnorm,1))/(tamano(1));
figure;plot(mediaC1L1); title('Espectro hueso Cordero 1')
```

Muestra C2L1 (L = lineal, C2 = Cordero 2)

```
cd ..
cd 03062025-113441_MedidaHueso_C2_L1_pumpenergy18
load ('matlabData.mat','spectra')
Espectro = spectra;
EspectroC2L1 = Espectro(:,7:15,:);
EspectroC2L1_2D = reshape(EspectroC2L1, size(EspectroC2L1,1), []);
EspectroC2L1_2D = EspectroC2L1_2D';
EspectroC2L1_2Dnorm = rescale(EspectroC2L1_2D, 0, 1);
tamano = size(EspectroC2L1_2D);
mediaC2L1 = (sum(EspectroC2L1_2Dnorm,1))/(tamano(1));
figure;plot(mediaC2L1); title('Espectro hueso Cordero 2')
```

Muestra CAL1 (L = lineal, CA = Cabra)

```
cd ..
cd 13062025-140048_MedidaHueso_CA_L1_pumpenergy25\
load ('matlabData.mat','spectra')
Espectro = spectra;
EspectroCAL1 = Espectro(:,7:15,:);
EspectroCAL1_2D = reshape(EspectroCAL1, size(EspectroCAL1,1), []);
EspectroCAL1_2D = EspectroCAL1_2D';
EspectroCAL1_2Dnorm = rescale(EspectroCAL1_2D, 0, 1);
tamano = size(EspectroCAL1_2Dnorm);
mediaCAL1 = (sum(EspectroCAL1_2Dnorm,1))/(tamano(1));
figure;plot(mediaCAL1); title('Espectro hueso Cabra medida lineal')
```

Muestra CAM1 (M = mapeo, CA = Cabra)

```
cd ..
cd 12062025-143201_MedidaHueso_CA_M1_pumpenergy25\
load ('matlabData.mat','spectra')
Espectro = spectra(:,7:15,:,:);
temp = permute(Espectro, [2, 3, 4, 1]); % Resultado: [9, 117, 24,
15049]
EspectroCAM1_2D = reshape(temp, [], size(Espectro, 1));
EspectroCAM1_2Dnorm = rescale(EspectroCAM1_2D, 0, 1);
tamano = size(EspectroCAM1_2Dnorm);
mediaCAM1 = (sum(EspectroCAM1_2Dnorm,1))/(tamano(1));
figure;plot(mediaCAM1); title('Espectro hueso Cabra medida con mapeo')
```

Muestra CR1M1 (M=Mapeo, CR1=Ciervo 1)

```
cd ..
cd 18072025-165646_MedidaHueso_CR1_M1_pumpenergy25_post\
load ('matlabData.mat','spectra')
Espectro = spectra(:,7:15,:,:);
temp = permute(Espectro, [2, 3, 4, 1]);
EspectroCR1M1_2D = reshape(temp, [], size(Espectro, 1));
EspectroCR1M1_2Dnorm = rescale(EspectroCR1M1_2D, 0, 1);
tamano = size(EspectroCR1M1_2Dnorm);
mediaCR1M1 = (sum(EspectroCR1M1_2Dnorm,1))/(tamano(1));
figure;plot(mediaCR1M1); title('Espectro hueso Ciervo 1 medida con
mapeo')
```

Muestra CR2M1 (M=Mapeo, CR2=Ciervo 2)

```
cd ..
cd 25062025-184153_MedidaHueso_CR2_M1_pumpenergy25\
load ('matlabData.mat','spectra')
Espectro = spectra(:,7:15,:,:);
temp = permute(Espectro, [2, 3, 4, 1]);
EspectroCR2M1_2D = reshape(temp, [], size(Espectro, 1));
EspectroCR2M1_2Dnorm = rescale(EspectroCR2M1_2D, 0, 1);
tamano = size(EspectroCR2M1_2D);
mediaCR2M1 = (sum(EspectroCR2M1_2Dnorm,1))/(tamano(1));
figure;plot(mediaCR2M1); title('Espectro hueso Ciervo 2 medida con
mapeo')
```

Separación de los datos en entrenamiento y test

```
EntrenamientoClase1 = EspectroC1L1_2Dnorm;
EntrenamientoClase2 = EspectroC2L1_2Dnorm;
EntrenamientoClase3 = [EspectroCAL1_2Dnorm;
EspectroCAM1_2Dnorm(1:10000,:)];
EntrenamientoClase4 = EspectroCR1M1_2Dnorm(1:5000,:);
```

```

EntrenamientoClase5 = EspectroCR2M1_2Dnorm(1:5000,:);
%Se seleccionan solo algunos datos en clase3, clase4 y clase5 ya que
sino son demasiados y el modelo sobreajusta
TestClase1 = EntrenamientoClase1(2952:3690,:);
EntrenamientoClase1 = EntrenamientoClase1(1:2951,:);
TestClase2 = EntrenamientoClase2(1330:1665,:);
EntrenamientoClase2 = EntrenamientoClase2(1:1329,:);
TestClase3 = EntrenamientoClase3(8001:10000,:);
EntrenamientoClase3 = EntrenamientoClase3(1:8000,:);
TestClase4 = EntrenamientoClase4(4001:5000,:);
EntrenamientoClase4 = EntrenamientoClase4(1:4000,:);
TestClase5 = EntrenamientoClase5(4001:5000,:);
EntrenamientoClase5 = EntrenamientoClase5(1:4000,:);

```

Creación de la matriz de entrenamiento (MatrizEntrenamiento)

```

MatrizEntrenamiento = [EntrenamientoClase1; EntrenamientoClase2;
EntrenamientoClase3; EntrenamientoClase4; EntrenamientoClase5];

```

Creación de la matriz de validación (Y)

```

filasC1L1 = size(EntrenamientoClase1);
filasC1L1 = filasC1L1(1);

filasC2L1 = size(EntrenamientoClase2);
filasC2L1 = filasC2L1(1);

filasC3L1 = size(EntrenamientoClase3);
filasC3L1 = filasC3L1(1);

filasC4L1 = size(EntrenamientoClase4);
filasC4L1 = filasC4L1(1);

filasC5L1 = size(EntrenamientoClase5);
filasC5L1 = filasC5L1(1);

Y = [ones(filasC1L1+filasC2L1,1); 2 * ones(filasC3L1,1); 3 *
ones((filasC4L1+filasC5L1),1)];
%Las clases 1 y 2 son dos tipos de muestras de la misma especie por lo
que son la misma categoria
%Las clases 4 y 5 son dos tipos de muestras de la misma especie por lo
que son la misma categoria
Ycat = categorical(Y);

```

Cálculo de error en el árbol de decisión con los datos de test

```

TestClases12 = [TestClase1; TestClase2];
TestClases45 = [TestClase4; TestClase5];
[yfit12,~] = FineTree.predictFcn(TestClases12);

```

```

[yfit3,~] = FineTree.predictFcn(TestClase3);
[yfit45,~] = FineTree.predictFcn(TestClases45);

count1 = sum(yfit12 == "2", 'all') + sum(yfit3 == "3", 'all'); %Errores
en la categoria 1
count2 = sum(yfit3 == "1", 'all') + sum(yfit3 == "3", 'all'); %Errores
en la categoria 2
count3 = sum(yfit45 == "1", 'all') + sum(yfit45 == "2", 'all');
%Errores en la categoria 3

fallos = count1 + count2 + count3; %Total de errores

ErrorFineTree =
(fallos*100)/(length(yfit12)+length(yfit3)+length(yfit45)); %Error del
1.4187%

```

Cálculo de error en el modelo SVM con los datos de test

```

[yfit12,~] = SVMmodel.predictFcn(TestClases12);
[yfit3,~] = SVMmodel.predictFcn(TestClase3);
[yfit45,~] = SVMmodel.predictFcn(TestClases45);

count1 = sum(yfit12 == "2", 'all') + sum(yfit3 == "3", 'all'); %Errores
en la categoria 1
count2 = sum(yfit3 == "1", 'all') + sum(yfit3 == "3", 'all'); %Errores
en la categoria 2
count3 = sum(yfit45 == "1", 'all') + sum(yfit45 == "2", 'all');
%Errores en la categoria 3

fallos = count1 + count2 + count3; %Total de errores

ErrorSVMmodel =
(fallos*100)/(length(yfit12)+length(yfit3)+length(yfit45)); %Error del
1.83%

```

Cálculo de error en el modelo de regresión logística con los datos de test

```

[yfit12,~] = LogRegression.predictFcn(TestClases12);
[yfit3,~] = LogRegression.predictFcn(TestClase3);
[yfit45,~] = LogRegression.predictFcn(TestClases45);

count1 = sum(yfit12 == "2", 'all') + sum(yfit3 == "3", 'all'); %Errores
en la categoria 1
count2 = sum(yfit3 == "1", 'all') + sum(yfit3 == "3", 'all'); %Errores
en la categoria 2
count3 = sum(yfit45 == "1", 'all') + sum(yfit45 == "2", 'all');
%Errores en la categoria 3

fallos = count1 + count2 + count3; %Total de errores

```

```
ErrorLogRegression =
(fallos*100)/(length(yfit12)+length(yfit3)+length(yfit45)); %Error del
3.51%
```

Muestra V1M1 (M=Mapeo, V1=Vértebra desconocida 1)

```
cd 19072025-124734_MedidaHueso_V1_M1_pumpenergy25_post\
load ('matlabData.mat','spectra')
load ('matlabData.mat','lambdas')
Espectro = spectra(:,7:15,,:);

temp = permute(Espectro, [2, 3, 4, 1]);

EspectroV1M1_2D = reshape(temp, [], size(Espectro, 1));

EspectroV1M1_2Dnorm = rescale(EspectroV1M1_2D, 0, 1);

tamano = size(EspectroV1M1_2Dnorm);
mediaV1M1 = (sum(EspectroV1M1_2Dnorm,1))/(tamano(1));
figure;plot(lambdas,mediaV1M1);
title('Espectro hueso desconocido vertebra medida con
mapeo');xlabel('longitud de onda [nm]');ylabel('Intensidad normalizada');
```

Muestra O1M1 (M=Mapeo, O1=Omoplato desconocido 1)

```
cd ..
cd 22072025-125301_MedidaHueso_O1_M1_pumpenergy25_post\
load ('matlabData.mat','spectra')
load ('matlabData.mat','lambdas')
Espectro = spectra(:,7:15,,:);

temp = permute(Espectro, [2, 3, 4, 1]);

EspectroO1M1_2D = reshape(temp, [], size(Espectro, 1));

EspectroO1M1_2Dnorm = rescale(EspectroO1M1_2D, 0, 1);

tamano = size(EspectroO1M1_2Dnorm);
mediaO1M1 = (sum(EspectroO1M1_2Dnorm,1))/(tamano(1));
figure;plot(lambdas,mediaO1M1);
title('Espectro hueso desconocido omoplato medido con
mapeo');xlabel('longitud de onda [nm]');ylabel('Intensidad normalizada');
```

Muestra Z1M1 (M=Mapeo, Z1=Muestra desconocida 1)

```
cd ..
cd 23072025-124843_MedidaHueso_Z1_M1_pumpenergy25_post\
load ('matlabData.mat','spectra')
```

```

load ('matlabData.mat','lambdas')
Espectro = spectra(:,7:15,,:);

temp = permute(Espectro, [2, 3, 4, 1]);

EspectroZ1M1_2D = reshape(temp, [], size(Espectro, 1));

EspectroZ1M1_2Dnorm = rescale(EspectroZ1M1_2D, 0, 1);

tamano = size(EspectroZ1M1_2Dnorm);
mediaZ1M1 = (sum(EspectroZ1M1_2Dnorm,1))/(tamano(1));
figure;plot(lambdas,mediaZ1M1);
title('Espectro hueso desconocido Z1 medido con mapeo');xlabel('longitud
de onda [nm]');ylabel('Intensidad normalizada');

```

Diferenciacion entre individuos de cordero

```

TestClase1 = EspectroC1L1_2Dnorm(2952:3690,:);
EntrenamientoClase1 = EspectroC1L1_2Dnorm(1:2951,:);

TestClase2 = EspectroC2L1_2Dnorm(1330:1665,:);
EntrenamientoClase2 = EspectroC2L1_2Dnorm(1:1329,:);

MatrizTrain = [EntrenamientoClase1; EntrenamientoClase2];
tamano1 = size(EntrenamientoClase1);
tamano2 = size(EntrenamientoClase2);
Y = [ones(tamano1(1),1); 2 * ones(tamano2(1),1)];
Ycat = categorical(Y);

[yfit1,~] = FineTreeIndividuos.predictFcn(TestClase1);
[yfit2,~] = FineTreeIndividuos.predictFcn(TestClase2);

count1 = sum(yfit1 == "2", 'all'); %Errores en la categoria 1
count2 = sum(yfit2 == "1", 'all'); %Errores en la categoria 2
fallos = count1 + count2;
porcentajeFallos = (fallos/(length(yfit1)+length(yfit2)))*100

```

Prueba clasificación muestras desconocidas V1,Z1,O1

```

load('EspectroV1M1_2Dnorm.mat')
load('EspectroZ1M1_2Dnorm.mat')
load('EspectroO1M1_2Dnorm.mat')

%Clase 1--> cordero; clase 2 --> cabra; clase 3 --> Ciervo rojo

%Muestra V1
[yfitV1,~] = FineTree.predictFcn(EspectroV1M1_2Dnorm);
countV1_cat1 = sum(yfitV1 == "1", 'all'); %Cuántas veces se predice
categoria 1

```

```

countV1_cat2 = sum(yfitV1 == "2", 'all'); %Cuantas veces se predice
categoria 2
countV1_cat3 = sum(yfitV1 == "3", 'all'); %Cuantas veces se predice
categoria 3
%Se predice cat1 19 veces, cat2 3128 y cat3 16853

%Muestra Z1
[yfitZ1,~] = FineTree.predictFcn(EspectroZ1M1_2Dnorm);
countZ1_cat1 = sum(yfitZ1 == "1", 'all');
countZ1_cat2 = sum(yfitZ1 == "2", 'all');
countZ1_cat3 = sum(yfitZ1 == "3", 'all');
%Se predice cat1 0 veces, cat2 3949 y cat3 16051

%Muestra 01
[yfit01,~] = FineTree.predictFcn(Espectro01M1_2Dnorm);
count01_cat1 = sum(yfit01 == "1", 'all')
count01_cat2 = sum(yfit01 == "2", 'all')
count01_cat3 = sum(yfit01 == "3", 'all')
%Se predice cat1 378 veces, cat2 3590 y cat3 16032

```

Referencias

- [1] **ROLDÁN, Alicia Marín**, et al. Laser-Induced breakdown spectroscopy (LIBS) for the analyses of faunal bones: assembling of individuals and elemental quantification. *Optik*, 2020, vol. 218, p. 164992. Disponible en: <https://www.sciencedirect.com/science/article/abs/pii/S0030402620308287>
- [2] **YOUSSEF, Doaa; MOTTO-ROS, Vincent; ABDEL-HARITH, Mohamed**. Machine learning-based novel approach of image analysis in LIBS elemental mapping for discriminating archeological human bone. *Spectrochimica Acta Part B: Atomic Spectroscopy*, 2025, vol. 224, p. 107-108. Disponible en: <https://www.sciencedirect.com/science/article/abs/pii/S0584854724002532>
- [3] **KASEM, M. A.**, et al. LIBS analysis of artificial calcified tissues matrices. *Talanta*, 2013, vol. 108, p. 53-58. Disponible en: <https://www.sciencedirect.com/science/article/abs/pii/S0584854714001104>
- [4] **Mirapeix, Jesús M.** Apuntes: La evolución histórica de la óptica y la fotónica (tema 1), p. 3-4.
- [5] National Geographic. Isaac Newton: el primer científico que explicó la naturaleza de los colores. Noelia Freire. Disponible en: https://www.nationalgeographic.com.es/ciencia/isaac-newton-primer-cientifico-explicar-naturaleza-colores_19587
- [6] **De Zátanyi Sándor**, (ifj.) Fized, CC BY-SA 3.0. Disponible en: <https://commons.wikimedia.org/w/index.php?curid=12226400>
- [7] RAE, espectro. <https://dle.rae.es/espectro>
- [8] <https://www.mundomicroscopio.com/longitud-de-onda/>
- [9] **Mirapeix, Jesús M.** Apuntes: La evolución histórica de la óptica y la fotónica (tema 1).
- [10] Espectro de Fraunhofer. Fuente: Wikipedia. Licencia: Dominio Público. Disponible en: https://es.wikipedia.org/wiki/L%C3%ADneas_de_Fraunhofer
- [11] **Mirapeix, Jesús M.** Apuntes: Midiendo el mundo a través de la luz (tema 5).
- [12] Nagwa. Emission and absorption spectra. Disponible en: <https://www.nagwa.com/en/explainers/469167813067/>
- [13] Triplenlace, Espectroscopía atómica (I): Tipos de espectros atómicos. Disponible en: <https://triplenlace.com/2013/01/01/espectro-continuo-de-emision-y-de-absorcion/>
- [14] **Ramos García, Ruben**. ¿Cómo se produce la luz?, 2015. Disponible en: <https://saberesyciencias.com.mx/2015/01/07/como-se-produce-la-luz/>
- [15] **Halpern, Joshua**. Espectros atómicos. Disponible en: [7.2: Espectros atómicos - LibreTexts Español](#).
- [16] **Rami Arieli**: "The Laser Adventure" Chapter 2.2 page 1. Versión en español por A. Requena, C.Cruz, A. Bastida y J. Zúñiga. Universidad de Murcia. Spain. Disponible en: [Rami Arieli: "The Laser Adventure"](#).

- [17] **Anabitarte García, Francisco; López Higuera, José M; Cobo García, Adolfo.** Tesis doctoral: Desarrollo de nuevos dispositivos y técnicas para espectroscopía inducida por láser.
- [18] GoPhotonics. What is Laser-Induced Breakdown Spectroscopy (LIBS)? Disponible en: [What is Laser-Induced Breakdown Spectroscopy \(LIBS\)? - GoPhotonics.com](https://www.gophotonics.com/what-is-laser-induced-breakdown-spectroscopy-lib/)
- [19] **Moreira Osorio, Lesther & Posada, Eduardo & Ponce, Luis.** (2012). Descubriendo la materia mediante luz láser. La espectroscopía de plasma inducido por láser (LIBS) como método para determinar la composición elemental de la materia. p 35-38.
- [20] Wikipedia, Nd:YAG láser. Disponible en: https://en.wikipedia.org/wiki/Nd:YAG_laser
- [21] **Alicia Marín Roldán, Vishal Dwivedi, José Yravedra Sainz de los Terreros, Pavel Veis.** Laser-Induced breakdown spectroscopy (LIBS) for the analyses of faunal bones: Assembling of individuals and elemental quantification, Optik, Volume 218, 2020. Disponible en: <https://www.sciencedirect.com/science/article/pii/S0030402620308287>.
- [22] Atomic spectra database. Disponible en: <https://www.nist.gov/pml/atomic-spectra-database>.
- [23] **DE CARVALHO, Gabriel Gustinelli Arantes, et al.** Recent advances in LIBS and XRF for the analysis of plants. *Journal of Analytical Atomic Spectrometry*, 2018, vol. 33, no 6, p. 919-944. Disponible en: <https://pubs.rsc.org/zh-tw/content/getauthorversionpdf/C7JA00293A>.
- [24] **Anabitarte García, Francisco; López Higuera, José M; Cobo García, Adolfo.** Tesis doctoral: Desarrollo de nuevos dispositivos y técnicas para espectroscopía inducida por láser, p. 118-126. Disponible en:
- [25] Instituto de catálisis y petroleoquímica. Espectroscopía de emisión atómica ICP-OES. Disponible en: https://icp.csic.es/archivos/uapoyo/2020/Descripcion%20ICP_02.pdf
- [26] **Anabitarte García, Francisco; López Higuera, José M; Cobo García, Adolfo.** Tesis doctoral: Desarrollo de nuevos dispositivos y técnicas para espectroscopía inducida por láser, p. 112-118.
- [27] Malvern Panalytical. Fluorescencia de rayos X (XRF). Disponible en: [Fluorescencia de rayos X: tecnología de XRF explicada | Malvern Panalytical](https://www.malvernpanalytical.com/es/Products/Fluorescence%20Spectroscopy/XRF).
- [28] **Mirapeix, Jesús M; Arniz-Mateos, Rosa; García-Escárgaza, Asier; Gutierrez-Zugasti, Igor; López-Higuera, José M; Cobo, Adolfo.** Virtual sampling: Archaeological implications of a new technique for elemental mapping of Mg/Ca ratios in marine mollusc shells.
- [29] **Díaz Noval, Ramón; Mirapeix Serrano, Jesús.** Trabajo fin de grado: Procesado de imágenes 2D LIBS para la detección de líneas de crecimiento en conchas de lapas.
- [30] **PANDHIJA, Shiwani, et al.** Contaminant concentration in environmental samples using LIBS and CF-LIBS. *Applied Physics B*, 2010, vol. 98, p. 231-241. Disponible en: <https://link.springer.com/article/10.1007/s00340-009-3763-x>
- [31] **POGGIALINI, Francesco, et al.** Catching up on calibration-free LIBS. *Journal of Analytical Atomic Spectrometry*, 2023, vol. 38, no 9, p. 1751-1771. Disponible en: <https://pubs.rsc.org/en/content/articlehtml/2023/ja/d3ja00130j>

- [32] **ZHANG, Ning, et al.** A brief review of calibration-free laser-induced breakdown spectroscopy. *Frontiers in Physics*, 2022, vol. 10, p. 887171. Disponible en: <https://www.frontiersin.org/journals/physics/articles/10.3389/fphy.2022.887171/full>
- [33] **SIOZOS, Panagiotis, et al.** Application of laser-induced breakdown spectroscopy and neural networks on archaeological human bones for the discrimination of distinct individuals. *Journal of Archaeological Science: Reports*, 2021, vol. 35, p. 102769. Disponible en: <https://www.sciencedirect.com/science/article/abs/pii/S2352409X20305605>
- [34] **CÁRDENAS-ESCUADERO, Jafet; GALÁN-MADRUGA, David; CÁCERES, Jorge O.** Laser-Induced Breakdown Spectroscopy as an Accurate Forensic Tool for Bone Classification and Individual Reassignment. *Applied Spectroscopy*, 2025, vol. 79, no 2, p. 241-259. Disponible en: <https://journals.sagepub.com/doi/epub/10.1177/00037028241277897>
- [35] **CÁCERES, Jorge O.; DE LOS TERREROS, José Yravedra Sainz.** A real-world approach to identifying animal bones and Lower Pleistocene fossils by laser induced breakdown spectroscopy. *Talanta*, 2021, vol. 235, p. 122780. Disponible en: <https://www.sciencedirect.com/science/article/abs/pii/S0039914021007013>
- [36] **IWAMURA, Edna Sadayo Miazato; SOARES-VIEIRA, José Arnaldo; MUÑOZ, Daniel Romero.** Human identification and analysis of DNA in bones. *Revista do Hospital das Clínicas*, 2004, vol. 59, p. 383-388. Disponible en: <https://www.scielo.br/j/rhc/a/g4jn3MBK8hxYSpYGP5vJGYw/>
- [37] **GERARDO, Cañete Betancourt.** Estimación de edad a partir del análisis del avance del canal medular de huesos largos en bovinos. En *Morfovirtual 2022*. Disponible en: <https://morfovirtual.sld.cu/index.php/morfovirtual22/2022/paper/viewPaper/140>
- [38] **SOTO, H.; CASTELLANOS, R.; TORIBIO, L. R.** Estudio métrico del canal medular del húmero como iniciador de la edad. *Estudios de Antropología Biológica*, 2012, vol. 5. <https://revistas.unam.mx/index.php/eab/article/download/34607/31541/81699>
- [39] **BUCKLEY, Michael, et al.** Species identification by analysis of bone collagen using matrix-assisted laser desorption/ionisation time-of-flight mass spectrometry. *Rapid Communications in Mass Spectrometry: An International Journal Devoted to the Rapid Dissemination of Up-to-the-Minute Research in Mass Spectrometry*, 2009, vol. 23, no 23, p. 3843-3854. Disponible en: [Species identification by analysis of bone collagen using matrix-assisted laser desorption/ionisation time-of-flight mass spectrometry - Buckley - 2009 - Rapid Communications in Mass Spectrometry - Wiley Online Library](https://onlinelibrary.wiley.com/doi/10.1002/rcm.4280)
- [40] **NEWMAN, Margaret E., et al.** Identification of archaeological animal bone by PCR/DNA analysis. *Journal of Archaeological Science*, 2002, vol. 29, no 1, p. 77-84. <https://www.sciencedirect.com/science/article/pii/S0305440301906880>
- [41] **MIRAPEIX, Jesús, et al.** A comparative analysis of elemental imaging of marine mollusc shells using Laser Induced Breakdown Spectroscopy. *Microchemical Journal*, 2025, p. 113756. Disponible en: <https://www.sciencedirect.com/science/article/pii/S0026265X25011105?via%3Dihub#b0245>

- [42] Microchipotle. Imagen sobre el aprendizaje supervisado, 2025. Disponible en: <https://microchipotle.com/aprendizaje-supervisado-en-machine-learning-definicion-funcionamiento-y-aplicaciones/>
- [43] **Calvo, Diego**. Imagen sobre el aprendizaje no supervisado, 2019. Disponible en: <https://www.diegocalvo.es/aprendizaje-no-supervisado/>
- [44] Momomarrero. Imagen sobre árbol de decision, 2022. Disponible en: <https://momomarrero.com/2022/02/10/el-arbol-de-decision-ejemplo-practico-i/>
- [45] Matlab Help Center: Classification Discriminant. Disponible en: <https://es.mathworks.com/help/stats/classificationdiscriminant.html>
- [46] Scikit learn: Linear and quadratic discriminant analysis, 2025. Disponible en: https://scikit-learn.org/stable/modules/lda_qda.html
- [47] **Ng, A.** (2017). *Supervised Machine Learning: Regression and Classification*. Disponible en: <https://www.coursera.org/learn/machine-learning>
- [48] **PÉREZ, JA Martínez; MARTÍN, PS Pérez**. Regresión logística. *Medicina de Familia. SEMERGEN*, 2024, vol. 50, no 1, p. 102086.
- [49] Matlab Help Center: Classification Emsembles. Disponible en: [Classification Ensembles - MATLAB & Simulink](#)
- [50] Dpto. de matemática aplicada, Universidad Politécnica de Madrid. Clasificación naive-Bayes, 2021. Disponible en: [02.1 Clasificación naive-Bayes — Introducción al Aprendizaje Automático](#)
- [51] Matlab Help Center: Support vector machine. Disponible en: <https://es.mathworks.com/discovery/support-vector-machine.html>
- [52] Matlab Help Center: Vecinos más próximos. Disponible en: [Vecinos más próximos - MATLAB & Simulink](#)
- [53] Scikit learn: Nearest neighbors. Disponible en: <https://scikit-learn.org/stable/modules/neighbors.html>
- [54] **Navarro, Sandra**. Características de los boosted trees, 2024. Disponible en: <https://keepcoding.io/blog/2-caracteristicas-de-los-boosted-trees/>
- [55] **Navarro, Sandra**. Boosted trees y la revolución del machine learning, 2024. Disponible en: [Boosted trees y la revolución del machine learning](#)
- [56] **Carmona, Francesc**. Bagged Trees, 2019. <https://www.ub.edu/cursosR/files/bagging.html>
- [57] **Andrew Ng**. DEEPLERNING.AI; COURSERA. *Advanced Learning Algorithms* [en línea]. Machine Learning Specialization. Coursera. Disponible en: <https://www.coursera.org/learn/advanced-learning-algorithms?specialization=machine-learning-introduction>
- [58] **Curo de la Cruz, Walther**. Esquema de red neuronal, 2024. Disponible en: <https://blog.walthercuro.com/crear-tu-primera-red-neuronal-artificial/>