

UNIVERSIDAD DE CANTABRIA

FACULTAD DE CIENCIAS

DEPARTAMENTO DE DE TECNOLOGÍA ELECTRÓNICA, INGENIERÍA
DE SISTEMAS Y AUTOMÁTICA



TRABAJO DE FIN DE GRADO

REDES NEURONALES Y DEEP LEARNING APLICADO A LA
CLASIFICACIÓN DE MATERIALES A PARTIR DE
ESPECTRO ÓPTICO Y ACÚSTICO

NEURAL NETWORKS AND DEEP LEARNING APPLIED TO
MATERIAL CLASSIFICATION FROM OPTICAL AND
ACOUSTIC SPECTRUM

Supervisores: Adolfo Cobo García y Jesus Mirapeix Serrano

Pedro Quilles Campo

Grado en Matemáticas, Mención en Informática

Curso académico 2021-2022

REDES NEURONALES Y DEEP LEARNING APLICADO A LA CLASIFICACIÓN DE MATERIALES A PARTIR DE ESPECTRO ÓPTICO Y ACÚSTICO

Resumen:

LIBS es una técnica utilizada para conocer la composición atómica de materiales en distintos estados. Esta técnica consiste en ablacionar el material y captar la información espectral emitida a cada longitud de onda mediante un espectrómetro óptico en el momento de emisión de plasma del material. Una aplicación habitual de la técnica LIBS es para identificar un material desconocido a partir de su espectro de emisión. Para ello, los modelos basados en redes neuronales destacan por su eficiencia y capacidad de ajuste. Este documento recoge un estudio sobre la efectividad de este tipo de algoritmos para la clasificación de ocho tipos de plásticos distintos. Además, se mostrará si el espectro acústico asociado a la generación explosiva del plasma contiene información relevante para la clasificación y tiene relevancia de cara a una clasificación.

Abstract:

LIBS is a technique used to know the atomic composition of materials in different states. This technique consists of ablating the material and capturing the spectral information emitted at each wavelength using an optical spectrometer at the time of plasma emission from the material. A common application of the LIBS technique is to identify an unknown material from its emission spectrum. For this purpose, models based on neural networks stand out for their efficiency and fitting capability. This paper presents a study on the effectiveness of this type of algorithms for the classification of eight different types of plastics. In addition, it will be shown whether the acoustic spectrum associated with the explosive generation of the plasma contains relevant information for classification and has relevance for classification.

Índice de figuras

1.	<i>Esquema del montaje experimental LIBS</i>	5
2.	<i>Ejemplo de espectro LIBS de una muestra de carbonato de calcio y magnesio, con las líneas de emisión de estos dos elementos resaltadas</i>	6
3.	<i>Algunas funciones de activación comunes</i>	7
4.	<i>Imagen de la universidad de Cantabria antes y después de aplicar filtros</i>	9
5.	<i>Proceso interno de aplicar un filtro sobre una imagen en su forma matricial.</i>	9
6.	<i>Distribución de las categorías en el dataset</i>	11
7.	<i>Espectro óptico laser LIBS, Plástico 1</i>	12
8.	<i>Espectro óptico laser LIBS, Plástico 2</i>	12
9.	<i>Espectro óptico laser LIBS, Plástico 3</i>	13
10.	<i>Espectro óptico laser LIBS, Plástico 4</i>	13
11.	<i>Espectro óptico laser LIBS, Plástico 5</i>	14
12.	<i>Espectro óptico laser LIBS, Plástico 6</i>	14
13.	<i>Espectro óptico laser LIBS, Plástico 7</i>	15
14.	<i>Espectro óptico laser LIBS, Plástico 8</i>	15
15.	<i>Muestra espectro acústico, Plástico 1</i>	16
16.	<i>Muestra espectro acústico, Plástico 2</i>	16
17.	<i>Muestra espectro acústico, Plástico 3</i>	17
18.	<i>Muestra espectro acústico, Plástico 4</i>	17
19.	<i>Muestra espectro acústico, Plástico 5</i>	18
20.	<i>Muestra espectro acústico, Plástico 6</i>	18
21.	<i>Muestra espectro acústico, Plástico 7</i>	19
22.	<i>Muestra espectro acústico, Plástico 8</i>	19
23.	<i>Varianza explicada frente a la n-ésima componente principal</i>	21
24.	<i>Proyección de los datos en el plano Componente 1 - Componente 2 (ACP)</i>	21
25.	<i>Arquitectura de la red entrenada para clasificación del espectro óptico</i>	23
26.	<i>Función de pérdida en los conjuntos de entrenamiento y validación en el conjunto obtenido por LIBS</i>	23
27.	<i>Precisión en los conjuntos de entrenamiento y validación de la red en el conjunto obtenido por LIBS.</i>	24
28.	<i>Matriz de confusión sobre el conjunto de test en el conjunto obtenido por LIBS</i>	24
29.	<i>Funcionamiento de un filtro lineal</i>	25
30.	<i>Arquitectura de la red entrenada para clasificación del espectro acústico</i>	26
31.	<i>Función de pérdida en los conjuntos de entrenamiento y validación de la red.</i>	27
32.	<i>Precisión en los conjuntos de entrenamiento y validación de la red en el conjunto acústico.</i>	27
33.	<i>Matriz de confusión sobre el conjunto de test en el conjunto acústico</i>	28

Índice

1. Introducción	4
2. Contexto	5
2.1. LIBS	5
2.1.1. Aplicaciones	6
2.2. Redes Neuronales y Deep Learning	7
2.2.1. Estructura del Algoritmo	7
2.2.2. Entrenamiento de Redes	8
2.2.3. Redes Neuronales Convolucionales	8
2.3. Análisis de Componentes Principales y LIBS	9
3. Clasificación de Materiales con LIBS	10
3.1. Introducción	10
3.2. Análisis exploratorio	10
3.2.1. Balanceo de la muestra	10
3.2.2. Similitudes entre los datos	11
3.2.3. Reducción de dimensionalidad - Espectro óptico	20
3.3. Clasificación	22
3.3.1. Clasificación a partir del espectro óptico	22
3.3.2. Clasificación a partir del espectro acústico	25
4. Conclusiones	29
5. Bibliografía	31

1. Introducción

Para averiguar la composición química de un material desconocido se han desarrollado infinidad de técnicas de análisis. Una de ellas, en la que se focaliza gran actividad investigadora en la actualidad, es la espectroscopía de ruptura inducida por láser (LIBS: Laser-induced breakdown spectroscopy), que consiste en el análisis de los espectros ópticos emitidos por un plasma, generado a su vez tras la ablación del material con un láser pulsado muy intenso.

Estos espectros contienen picos de emisión muy bien definidos de los diferentes elementos químicos presentes en el material, que se identifican mediante su longitud de onda, diferente para cada elemento[1]. La interpretación de los espectros para averiguar de qué material se trata involucra normalmente la intervención de un experto humano, pero las aplicaciones de monitorizado en tiempo real se podrían beneficiar de un análisis automático.

Algún ejemplo de estas aplicaciones es el control de calidad de aceros a partir de las concentraciones de sus diferentes elementos químicos, la clasificación automática de chatarra metálica o plásticos[3] en plantas de reciclado, la detección de contaminantes o sustancias tóxicas en productos alimentarios, etc.

En este trabajo, se explora la posibilidad de clasificar automáticamente materiales desconocidos a partir de su espectro LIBS, utilizando algoritmos de aprendizaje automático, en particular, redes neuronales multicapa, con técnicas adicionales de pre-procesado y de reducción de dimensionalidad como PCA. Adicionalmente, se pretende averiguar si la capacidad de clasificación puede mejorar usando como información, además del espectro óptico, la información acústica generada por las ondas de choque asociadas a la producción del plasma tras la ablación.

En este trabajo se utilizan las medidas de espectros ópticos LIBS y acústicos de 8 tipos de plásticos de composición desconocida, que sirven para el entrenamiento de los modelos.

Tras esta introducción, el capítulo 2 analiza el contexto asociado a este trabajo: la espectroscopía LIBS y una breve introducción a distintos algoritmos de deep learning que se utilizan durante el trabajo realizado.

En el capítulo 3 se considera el cuerpo del trabajo, la metodología y los procedimientos utilizados, considerando las arquitecturas finales elegidas para la clasificación basada para el espectro óptico y acústico. Esto se inicia mediante un análisis exploratorio de los datos y finaliza con ambas clasificaciones. Finalmente, el capítulo 4 recoge las conclusiones y otras líneas con las que se podría continuar este trabajo.

2. Contexto

2.1. LIBS

LIBS (Laser Induced Breakdown Spectroscopy) es una técnica de espectroscopía de emisión atómica. A pesar de estar desarrollada previamente, su desarrollo se vio impulsado a partir de la década de los 80 debido al desarrollo, abaratamiento y disponibilidad de láseres y detectores ópticos, lo que causó una renovación del interés por la técnica.

Como síntesis, para poder reproducir esta técnica es necesario un láser, un espectrómetro, y un generador de pulsos que sincronice el disparo del láser con la captura de los espectros de emisión del plasma. El láser es de tipo pulsado y gran energía por pulso, siendo capaz de ablacionar el material, produciéndose una pequeña cantidad de plasma que emite la luz captada por el espectrómetro. Una vez captada la luz, se almacena la intensidad obtenida por cada longitud de onda.

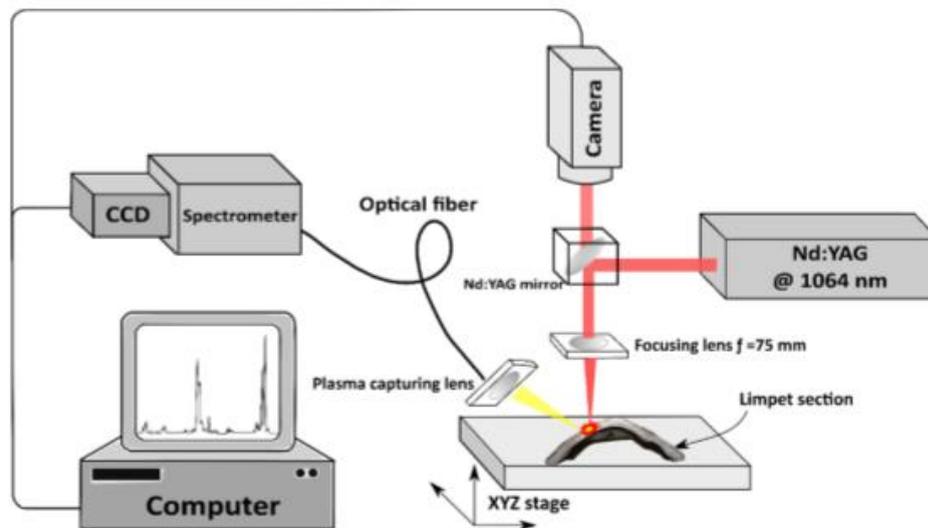


Figura 1: *Esquema del montaje experimental LIBS*

A su vez, cada longitud de onda puede representar la presencia de un material. Por ejemplo, las longitudes de onda 404(nm) y 437(nm) son características del hierro, mientras que las longitudes 514(nm) y 521(nm) son las generadas por los átomos de cobre. En cada espectro de emisión están presentes, típicamente, al menos varias líneas de emisión de cada elemento químico: su longitud de onda e intensidad ayudan a identificar el material ablacionado.

En la figura que se observa a continuación, se resaltan las longitudes de onda relevantes para calcio y magnesio. Esto puede ser usado con distintos fines que ya se están investigando [2].

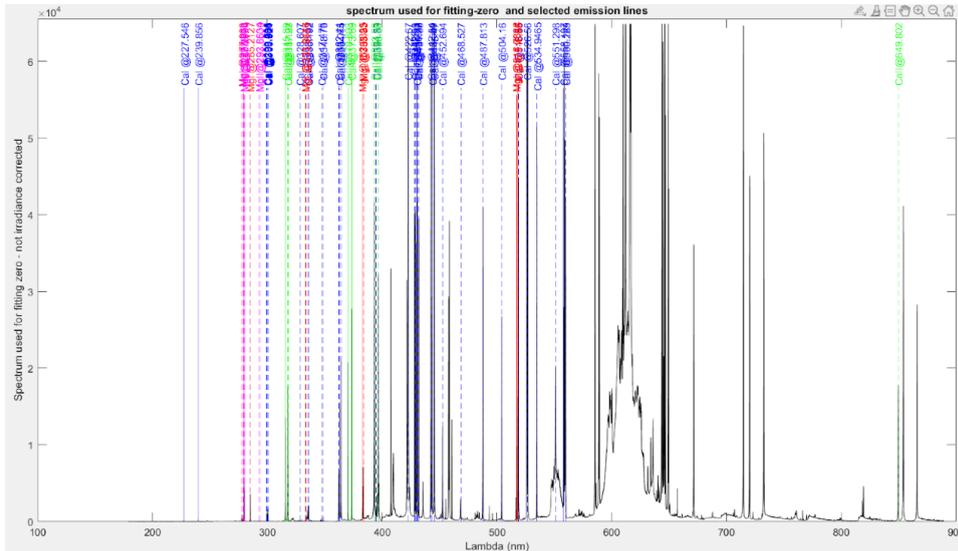


Figura 2: *Ejemplo de espectro LIBS de una muestra de carbonato de calcio y magnesio, con las líneas de emisión de estos dos elementos resaltadas*

2.1.1. Aplicaciones

La técnica LIBS permite determinar la presencia de ciertos elementos en el material analizado. También permite estimar cantidades y concentraciones. Esta capacidad hace que LIBS tenga aplicaciones en ámbitos diversos.

LIBS, además, se puede considerar una técnica casi no destructiva, ya que el material ablacionado puede ser tan poco como unos picogramos dejando un cráter de unas pocas micras. Esto permite su aplicación en ámbitos especialmente delicados, como joyería, o arte. Un tipo de aplicación común es la identificación de los tipos de pigmentos utilizados en pinturas[7].

Además, por motivos como estos, esta técnica se está haciendo cada vez más popular, aumentando el número de participantes en conferencias que abordan este tema y aumentando el número de artículos publicados año a año [8].

Otro ejemplo destacable de las aplicaciones de esta técnica es en el ámbito biomédico. El uso de esta técnica en materiales orgánicos duros han permitido completar investigaciones relacionando la dureza de los huesos con magnesio y calcio, y rápida detección de caries en dientes por las emisiones de calcio, magnesio, cobre, zinc, titanio, fósforo, potasio, hidrógeno, carbono y oxígeno.

Para su aplicación en técnicas como estas, la clasificación de materiales con técnicas de machine learning toman un papel importante, permitiendo reducir la complejidad de los problemas y agilizarla en caso necesario.

2.2. Redes Neuronales y Deep Learning

2.2.1. Estructura del Algoritmo

Una red neuronal artificial común se compone por perceptrones (las neuronas) y las capas. A su vez, cada perceptrón es una composición de funciones con N entradas y una salida. Los perceptrones están divididos a su vez en dos partes, la transformación lineal y la función de activación. La transformación lineal actúa a partir de los pesos (w), que son coeficientes reales que multiplican la entrada de la siguiente manera:

$$f(\bar{x}) = W^T x + b = w_1 * x_1 + \dots + w_n * x_n + b \quad w_i, b \in R, i \in \{1, \dots, n\}$$

El resultado de la transformación lineal es frecuentemente denotado por z . Una vez calculado z , es necesario componer el resultado con una función no lineal denominada función de activación. La motivación es para romper la linealidad de la red. De no hacerse, añadir más capas no aportaría ningún cambio, por lo que al final se seguiría tratando de una clasificación lineal. Además, el añadir más capas permite crear una arquitectura más compleja que pueda amoldarse mejor a los datos, resultando en una predicción más acertada o una clasificación más precisa, por lo que no es algo sacrificable. El término deep learning es utilizado para representar las redes neuronales profundas, denominadas así para aquellas que están compuestas por más de 4 o 5 capas. Algunas de las activaciones más populares son la tangente hiperbólica, la función sigmoide y la denominada ReLU:

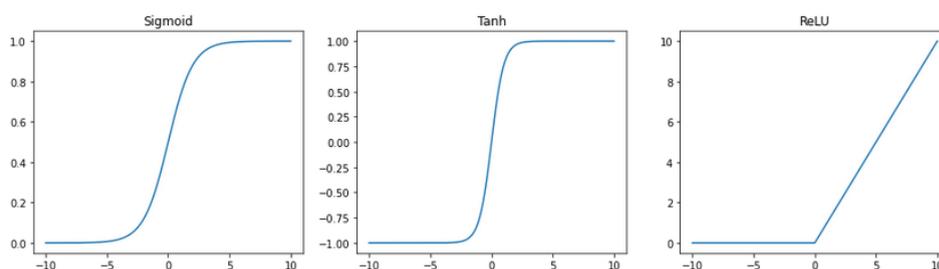


Figura 3: Algunas funciones de activación comunes

La función sigmoide es una transformación de la tangente hiperbólica. De hecho $sigmoid(x) = \frac{tanh(\frac{x}{2}) + \frac{1}{2}}{2}$. La diferencia es que la función sigmoide es más suave que la tangente hiperbólica. La mayor diferencia reside en la ReLU. Esta activación ayuda a resolver el problema de desvanecimiento de gradiente de las redes neuronales. Por este motivo, esta será la activación utilizada en las redes que consideraremos.

El desvanecimiento de gradiente (vanishing gradients en inglés) es un problema que causa un aumento excesivo en el tiempo de entrenamiento en las redes neuronales con muchas capas, que ha sido mitigado eficazmente mediante el descubrimiento de este tipo de activaciones [6].

Finalmente, en cada capa de la red, pueden situarse tantos perceptrones como se considere apropiado. La dimensión de la entrada de cada perceptrón será igual al número de variables de nuestro problema en la primera capa, e igual al número de perceptrones de la capa anterior para todos los demás casos. Con esta construcción, la ejecución de una red neuronal se inicia en las primeras capas y se continúa en las siguientes

a medida que se van obteniendo los resultados de cada capa. Esto es conocido como forward propagation.

2.2.2. Entrenamiento de Redes

El término “entrenar” representa el proceso de ajuste de los pesos propios de cada perceptrón. Las redes neuronales tienen la capacidad de replicar cualquier función existente (universal approximation theorems) siempre que tengan un número de capas y neuronas suficiente. Los pesos de cada perceptrón son inicializados aleatoriamente, por lo que necesitan ajustarse de manera que produzcan la salida adecuada.

Para entrenar una red se utilizan funciones de pérdida (loss functions) para estimar cómo de equivocada está la red y cuánto tiene que ajustarse. Vamos a tomar como ejemplo la función de pérdida para clasificación binaria, con una red con una salida de dimensión uno. En este caso, dado un input, nuestra red nos daría un número entre 0 y 1 que es la probabilidad de pertenecer (o no) a una determinada clase. Para este caso, la función de pérdida será la siguiente:

$$C(w, b) = \frac{1}{m} \sum_{i=1}^m \mathcal{L}(y_{real}^{(i)}, y_{pred}^{(i)}) = -\frac{1}{m} \sum_{i=1}^m (y_{real}^{(i)} \log(y_{pred}^{(i)}) + (1 - y_{real}^{(i)}) \log(1 - y_{pred}^{(i)})) \quad (1)$$

El ajuste de los pesos del perceptrón se realiza mediante el descenso del gradiente. Esto es, cada w pasa a ser $w - \alpha \frac{\partial C(w)}{\partial w}$ siendo alpha un número real denominado *learning rate*. De esta manera, los pesos se ajustan ‘mágicamente’ de manera que la red produce la salida deseada.

2.2.3. Redes Neuronales Convolucionales

Las redes neuronales convolucionales funcionan de una manera algo distinta. Sin embargo, su comportamiento es aún más interesante. Este tipo de redes tiene como objetivo el procesamiento de datos multidimensionales, como imágenes o espacios. Este tipo de redes se dividen por capas que contienen filtros.

Los filtros se llevan utilizando durante más de 60 años con el objetivo de extraer información de las imágenes (Entendidas como 3 matrices concatenadas, habitualmente representando su intensidad de color en rojo, verde y azul.). Al aplicar un filtro a una imagen, se obtienen distintos resultados. Por ejemplo, se pueden utilizar filtros distintos para obtener los bordes de una imagen, sean verticales, horizontales o generales.

De manera resumida, un filtro reúne las características de una zona (la parte de la imagen sobre la que se aplica) según unos criterios específicos (los valores del filtro). Los valores se calculan mediante la convolución de matrices:

$$\begin{bmatrix} x_{11} & x_{12} & \cdots & x_{1n} \\ x_{21} & x_{22} & \cdots & x_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ x_{m1} & x_{m2} & \cdots & x_{mn} \end{bmatrix} * \begin{bmatrix} y_{11} & y_{12} & \cdots & y_{1n} \\ y_{21} & y_{22} & \cdots & y_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ y_{m1} & y_{m2} & \cdots & y_{mn} \end{bmatrix} = \sum_{i=0}^{m-1} \sum_{j=0}^{n-1} x_{(m-i)(n-j)} y_{(1+i)(1+j)}$$

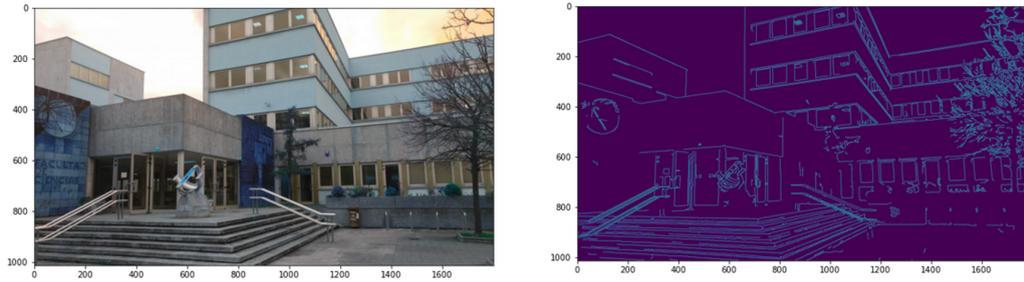


Figura 4: Imagen de la universidad de Cantabria antes y después de aplicar filtros

Al diseñar la arquitectura de una red convolucional, algunos de los parámetros de estas redes son el número de filtros y el número de filas y columnas de cada filtro. Al entrenar una red convolucional bien diseñada, cada filtro se ajusta de una manera distinta. Las primeras capas suelen recoger las formas básicas de cada imagen, rectas verticales, horizontales, curvas, etc. Las siguientes capas, se ajustan para detectar conjuntos de formas, incluyendo cuadrados, círculos y otras formas características que ayudan a la clasificación. A medida que cada filtro es entrenado, van cambiando sus valores internos, lo que produce una salida distinta.

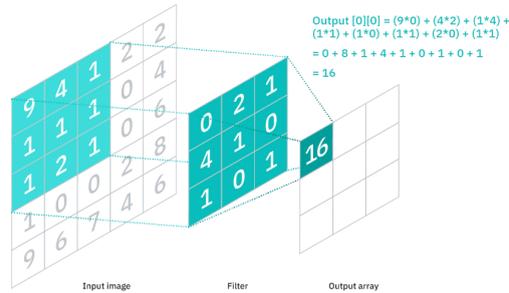


Figura 5: Proceso interno de aplicar un filtro sobre una imagen en su forma matricial.

2.3. Análisis de Componentes Principales y LIBS

Para procesar el espectro óptico, se ha optado por utilizar un análisis de componentes principales.

A parte de ser un excelente método de reducción de dimensionalidad, el análisis de componentes principales se ha utilizado junto a LIBS como una herramienta efectiva para el análisis exploratorio y el reconocimiento de patrones, generando una reducción de dimensionalidad basada en la varianza del conjunto original[10].

Por ser un método efectivo, se ha decidido conscientemente utilizarse a pesar de no cumplir las hipótesis necesarias para el ACP que son: (1) Las variables están linealmente correladas y (2) su distribución es una distribución normal[8]. A pesar de ello, el método ACP es ampliamente utilizado por la comunidad de investigadores [10].

3. Clasificación de Materiales con LIBS

3.1. Introducción

El problema tratado es un problema de clasificación entre ocho tipos de plásticos distintos. La información que tenemos de cada muestra es un conjunto de 15211 variables correspondientes a la intensidad recogida en cada una de las 15211 longitudes de onda captadas. Además, para cada muestra se tiene también un conjunto de 100 variables que corresponden a la serie temporal de la intensidad del sonido en el momento del disparo. Por el momento, no se tiene constancia de si el sonido contiene información útil para la clasificación. Además, clasificar un material según su espectro óptico no es sencillo, por lo que otra línea de desarrollo es investigar si ambos grupos de variables en conjunto colaboran de cara a una mejor clasificación.

3.2. Análisis exploratorio

En el análisis exploratorio se analizan distintas cuestiones, como el balanceo del dataset, el tamaño de la muestra, la similitud del dataset. Y unas muestras visuales sobre el tipo de datos que se están observando.

3.2.1. Balanceo de la muestra

El balanceo de una muestra es un factor clave a estudiar. Una gran parte de los problemas de clasificación existentes cuya resolución tiene interés, al menos económico, no están balanceados. Esto da lugar a problemas en el entrenamiento de modelos, resultando en clasificadores que atribuyen toda la muestra como perteneciente a la clase mayoritaria, obteniendo así una precisión (Entendida como el número de muestras correctamente clasificadas con respecto al total) igual a la proporción del conjunto mayoritario con respecto al total.

Para contrarrestar este problema, es común el uso de técnicas de balanceo como el oversampling, que consiste en añadir muestras iguales del conjunto minoritario al conjunto de entrenamiento. Esto aumenta la probabilidad de overfitting, esto es, que el modelo aprenda del ruido de la muestra (distintos factores aleatorios comunes en el conjunto de entrenamiento) en vez de en las propiedades intrínsecas de la muestra, en este caso, de los plásticos.

Por otro lado, se podría considerar el undersampling, que consiste en eliminar muestras de las clases mayoritarias con el objetivo de que estén balanceadas. Sin embargo, en este caso se elimina información que podría ser relevante para la clasificación.

En este caso específico, el dataset que es objeto de estudio tiene un total de 1209 muestras. Como se puede observar en la figura inferior, las categorías están bien distribuidas, por lo que no se ha considerado necesario realizar ninguna técnica de balanceo de variables como los mencionados anteriormente.

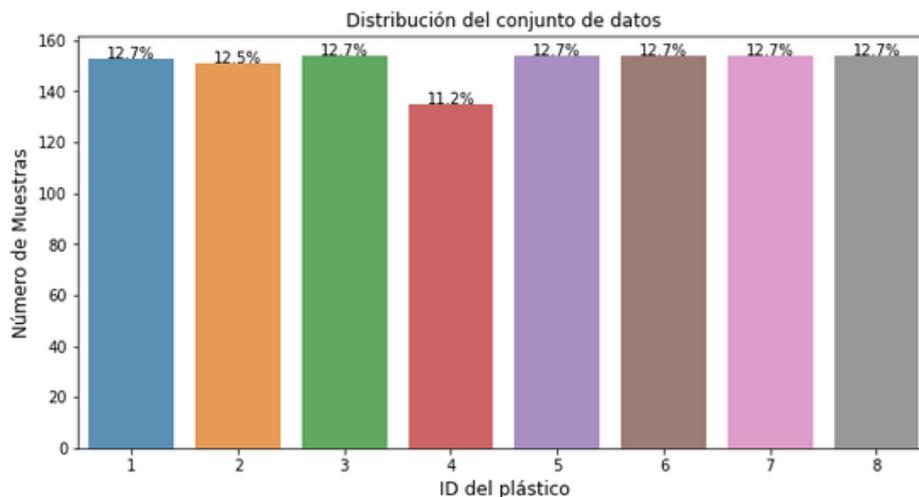


Figura 6: *Distribución de las categorías en el dataset*

ID del plástico	1	2	3	4	5	6	7	8
Número de Muestras	153	151	154	135	154	154	154	154

3.2.2. Similitudes entre los datos

El espectro LIBS tiene el inconveniente de no ser una técnica exacta. La luz emitida por el material que ha sido impactado por el laser se origina a partir del plasma. Esto implica que cualquier variación de temperatura, humedad y otro sinnúmero de factores difícilmente identificables y replicables pueden afectar a los valores obtenidos. Aún en el caso de realizar las tomas en momentos consecutivos sobre el mismo material, los valores difieren entre sí.

Por este motivo, vamos a mostrar para cada uno de los plásticos, cuatro muestras aleatorias del espectro. Las figuras aparecerán a continuación. Bajo los 4 gráficos, se encontrará un pequeño comentario indicando al grupo al que pertenecen.

Cabe destacar que en muchos casos, aunque la forma pudiera ser parecida, vemos que las intensidades son mucho mayores. Esto se debe a que al aplicar la técnica LIBS, se debe seleccionar la intensidad del laser. Si la intensidad fuera muy reducida, el espectro estaría formado en gran parte de ruido, obtenido algo parecido a lo que se puede observar en dos casos de la Figura 9, que expone dos muestras del espectro óptico característico del plástico 4. Por otro lado, si se elevara la intensidad, a medida que las longitudes de onda que con mayor respuesta aumenta, las intensidades correspondientes a las longitudes de onda más débiles serían aún menores, hasta permanecer indistinguibles con otras que no aportan nada.

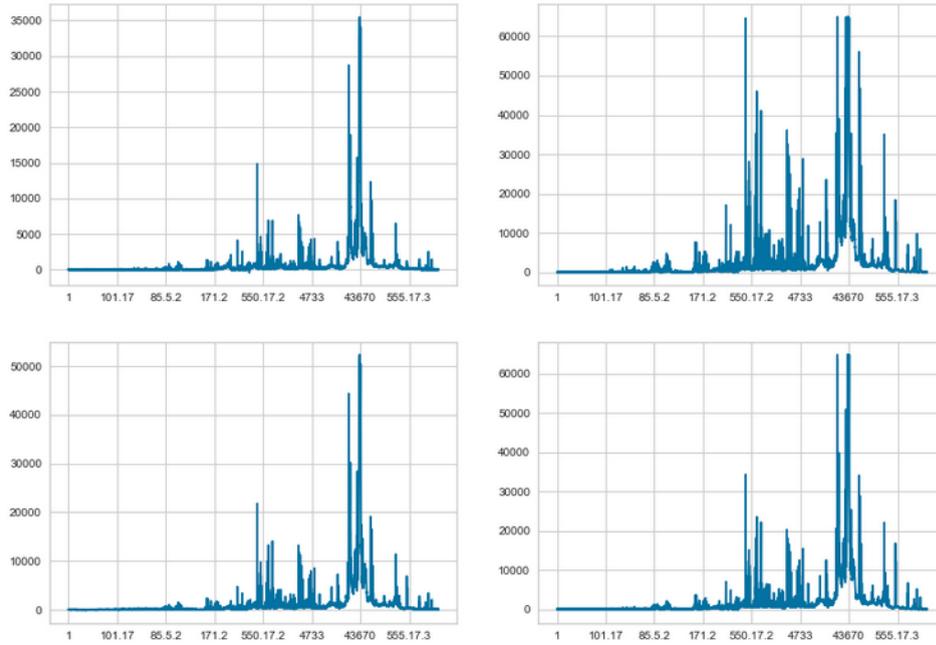


Figura 7: Espectro óptico laser LIBS, Plástico 1

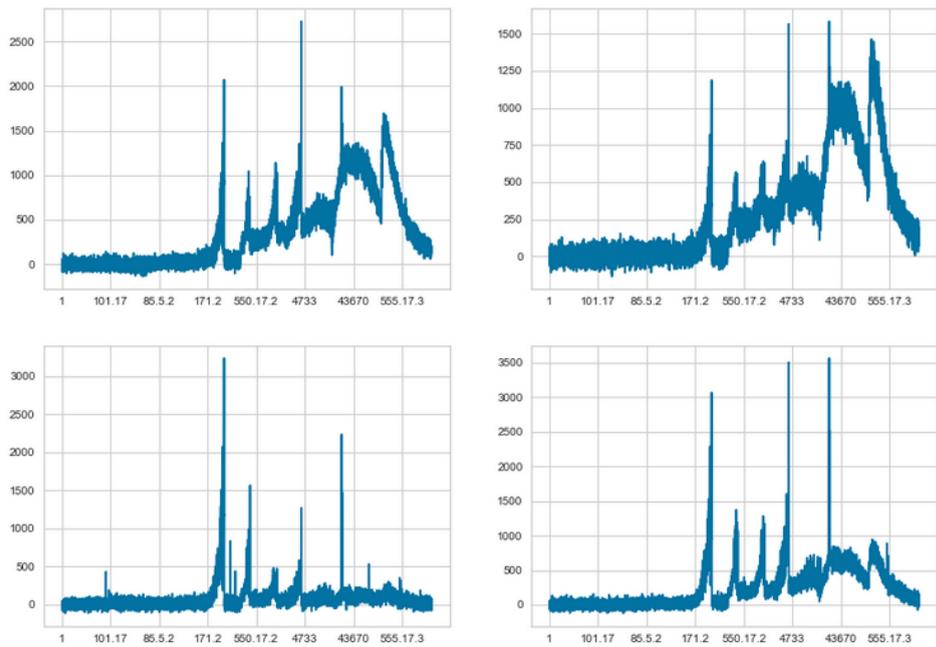


Figura 8: Espectro óptico laser LIBS, Plástico 2

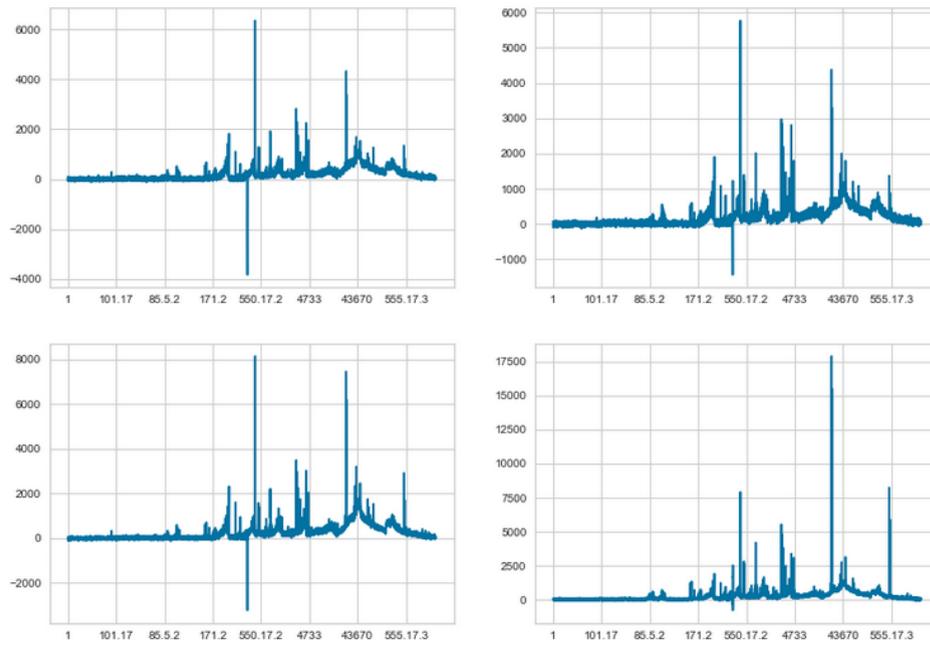


Figura 9: *Espectro óptico laser LIBS, Plástico 3*

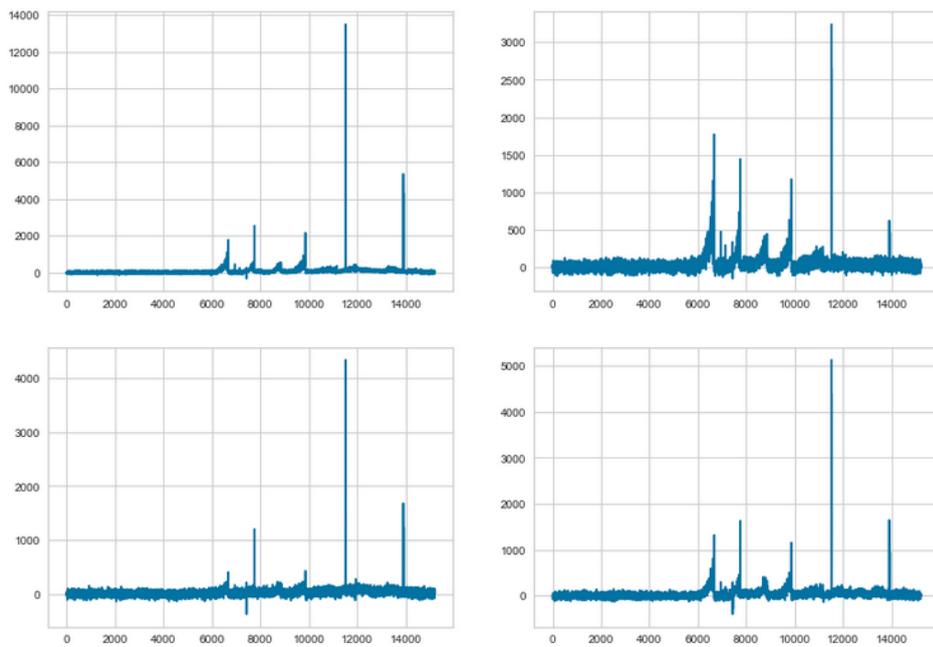


Figura 10: *Espectro óptico laser LIBS, Plástico 4*

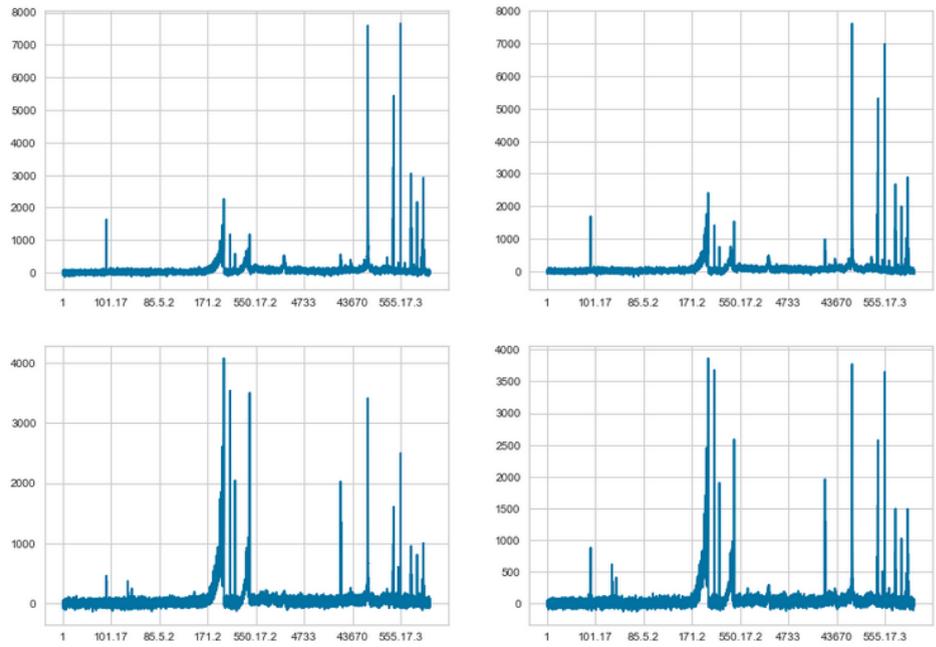


Figura 11: *Espectro óptico laser LIBS, Plástico 5*

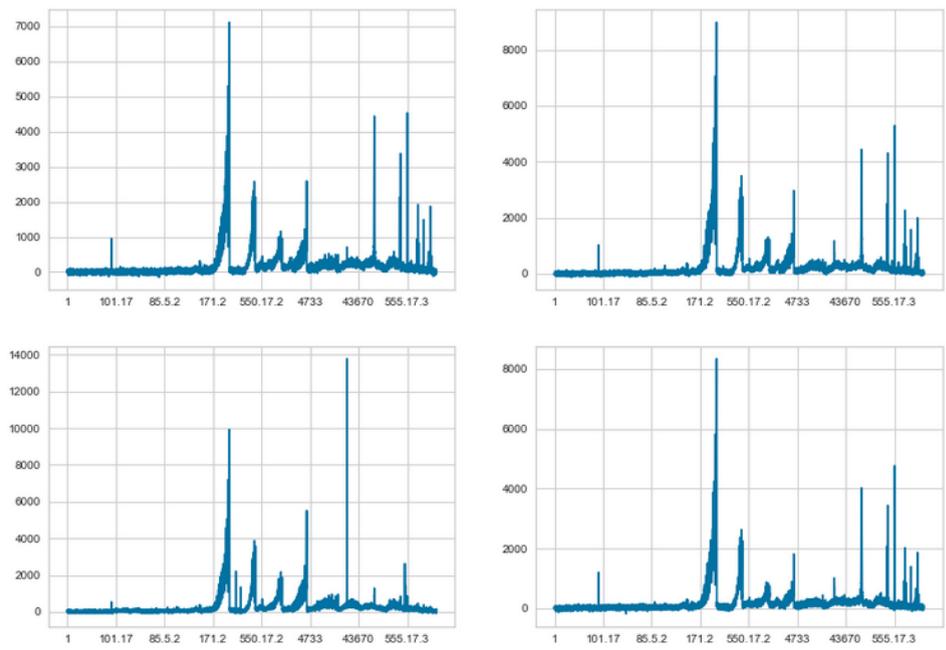


Figura 12: *Espectro óptico laser LIBS, Plástico 6*

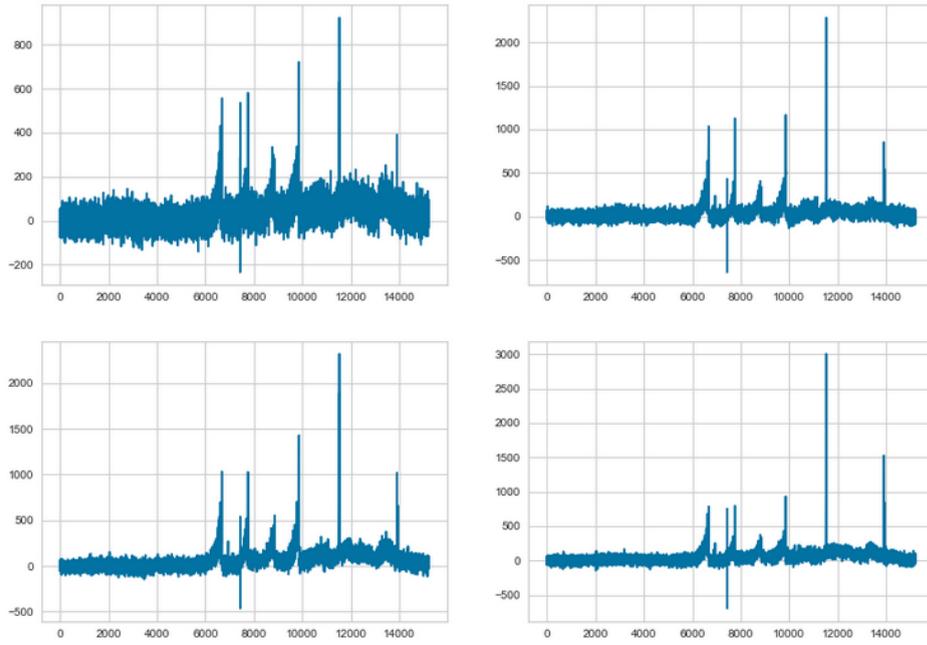


Figura 13: *Espectro óptico laser LIBS, Plástico 7*

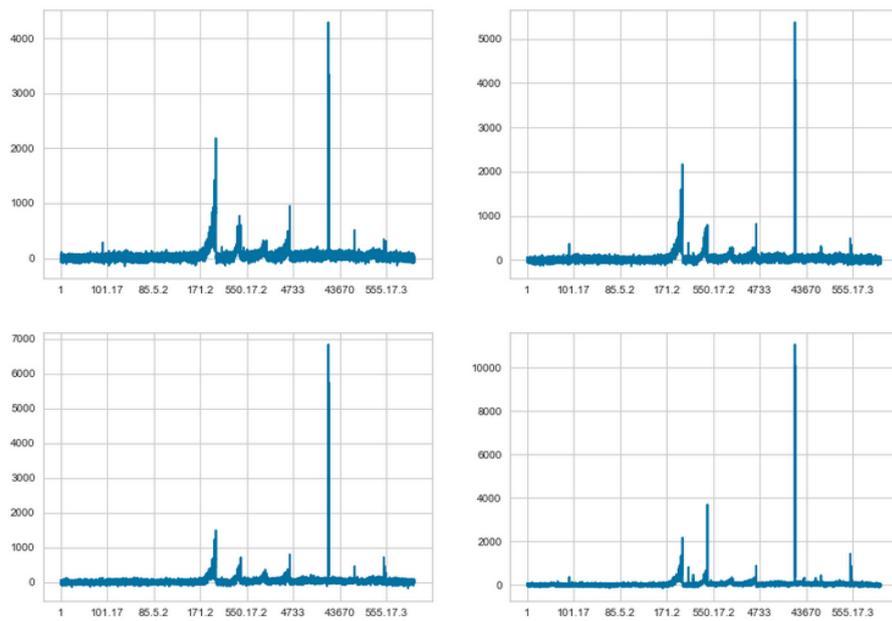


Figura 14: *Espectro óptico laser LIBS, Plástico 8*

Por otro lado, mostraremos la serie temporal de la intensidad del sonido en el momento de la toma de la muestra para cada tipo de plástico. Los espectros ópticos mostrados anteriormente corresponden a a misma muestra de la cual se mostrarán los ejemplos de los espectros acústicos según posición y plástico.

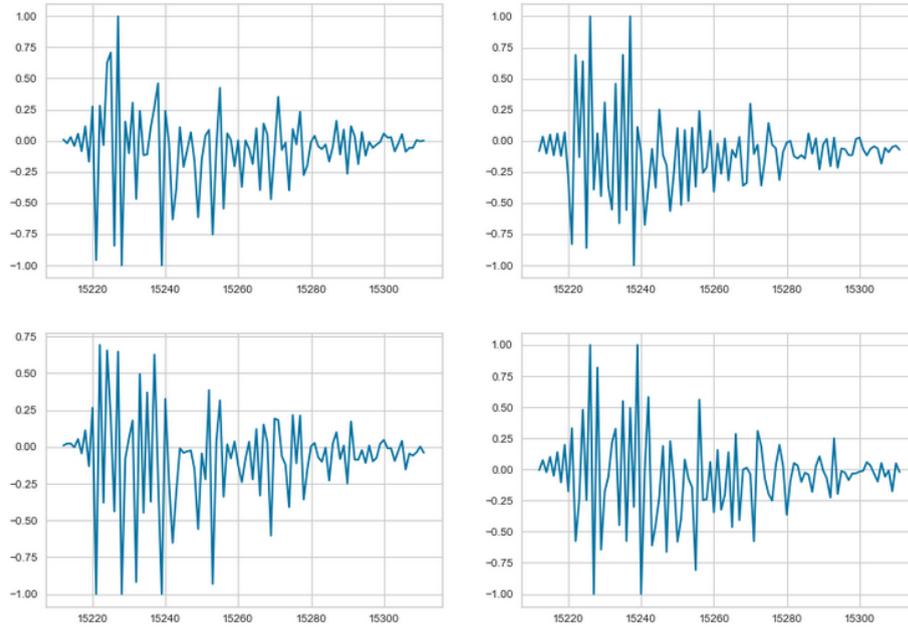


Figura 15: *Muestra espectro acústico, Plástico 1*

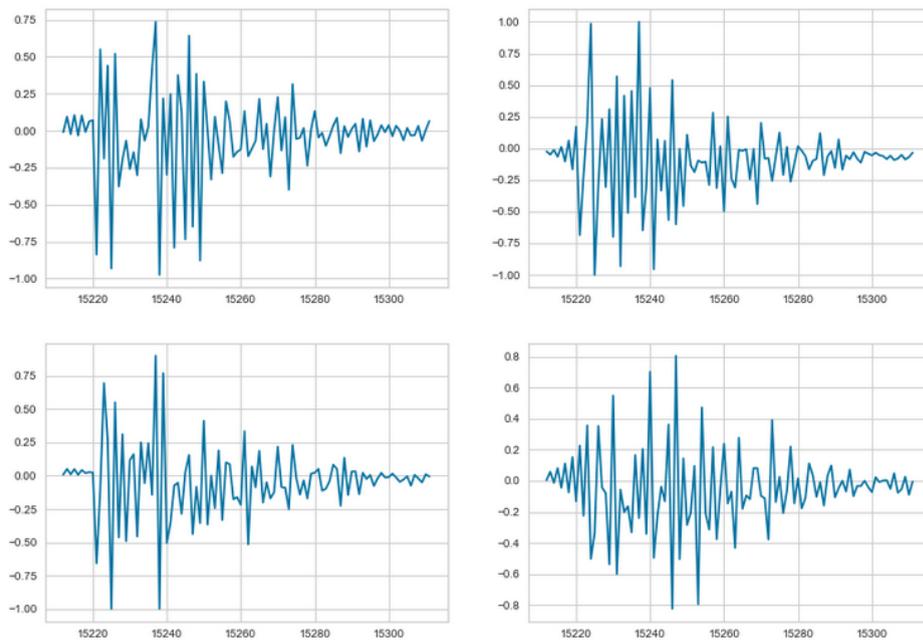


Figura 16: *Muestra espectro acústico, Plástico 2*

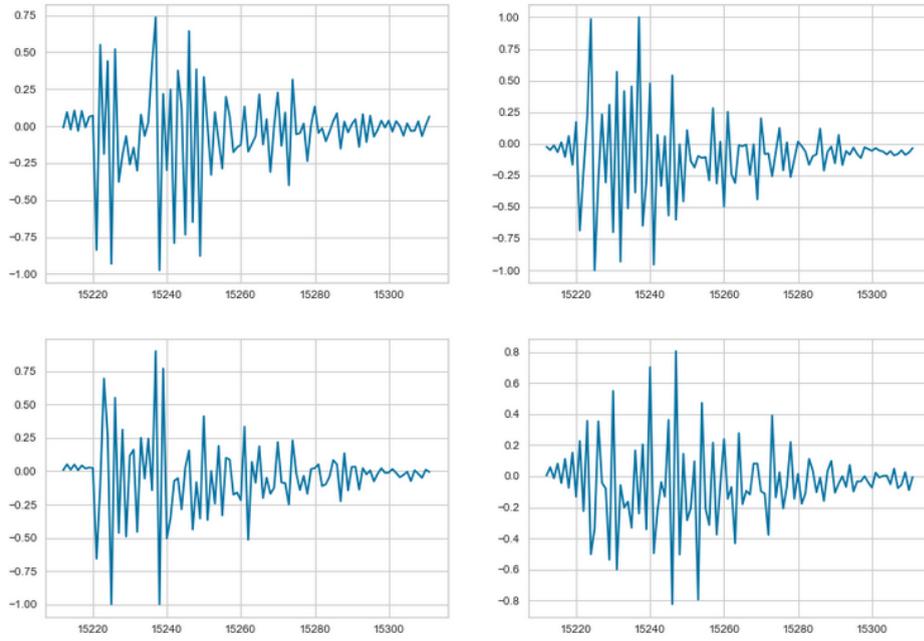


Figura 17: *Muestra espectro acústico, Plástico 3*

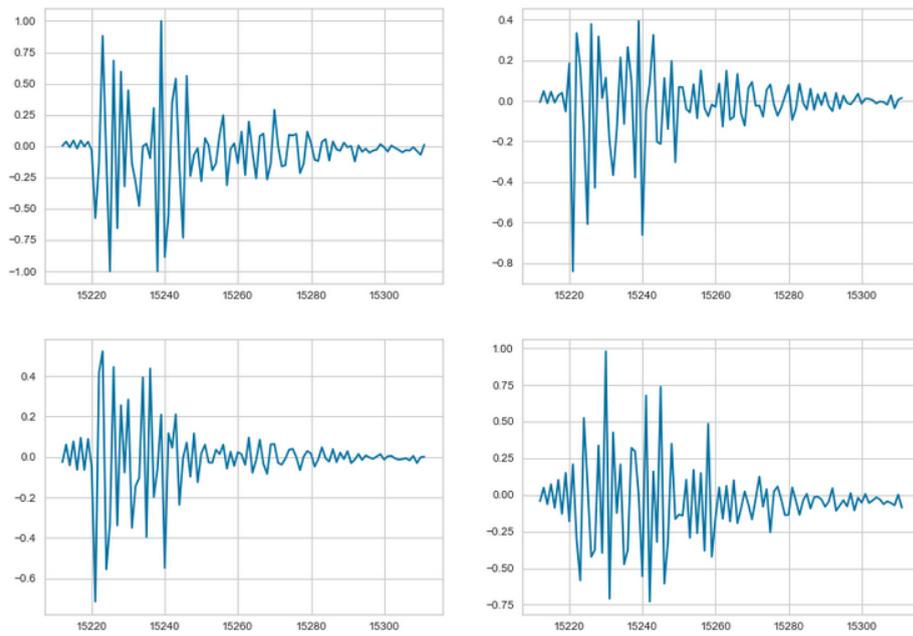


Figura 18: *Muestra espectro acústico, Plástico 4*

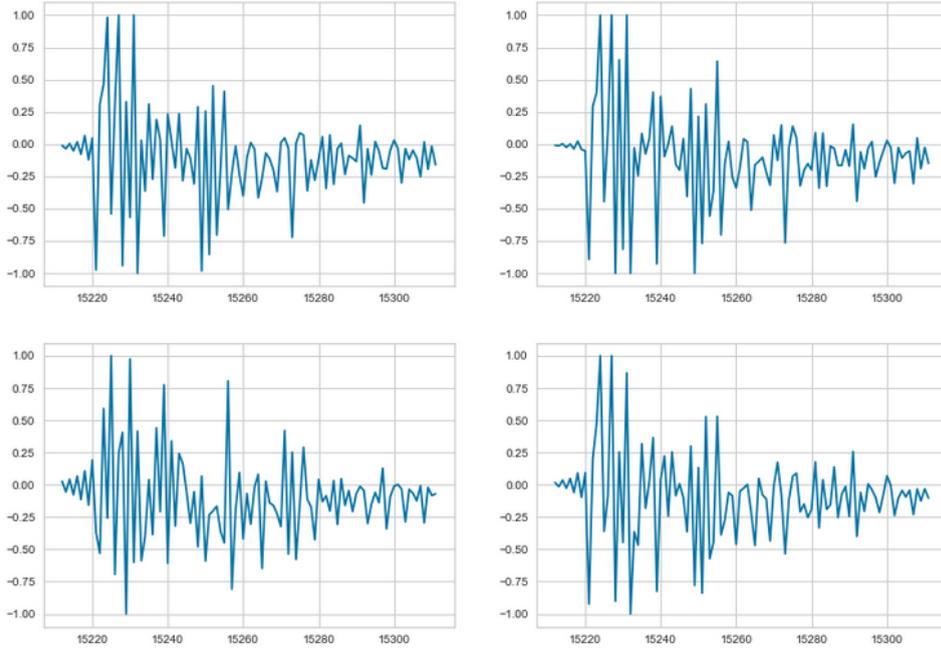


Figura 19: *Muestra espectro acústico, Plástico 5*

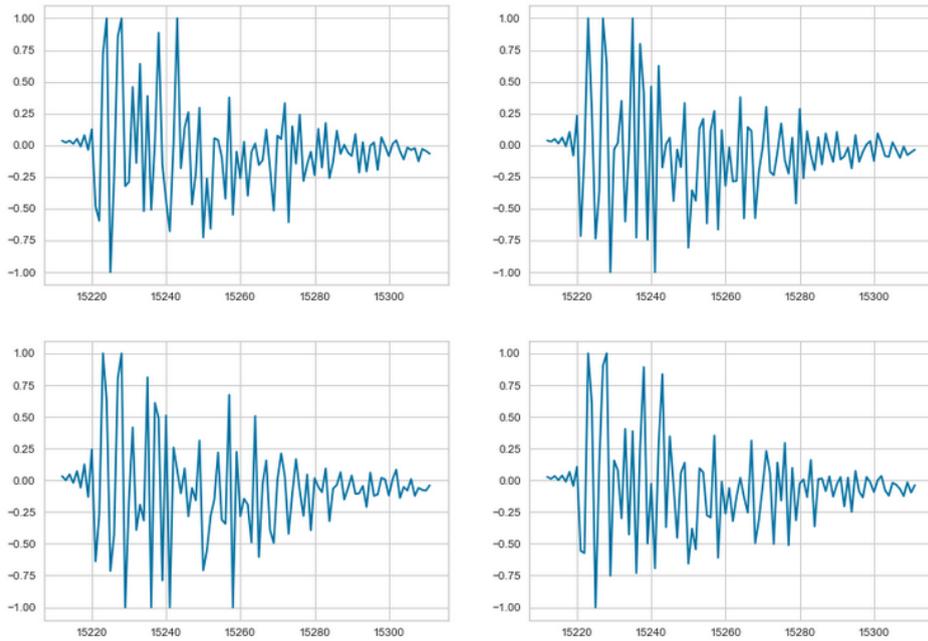


Figura 20: *Muestra espectro acústico, Plástico 6*

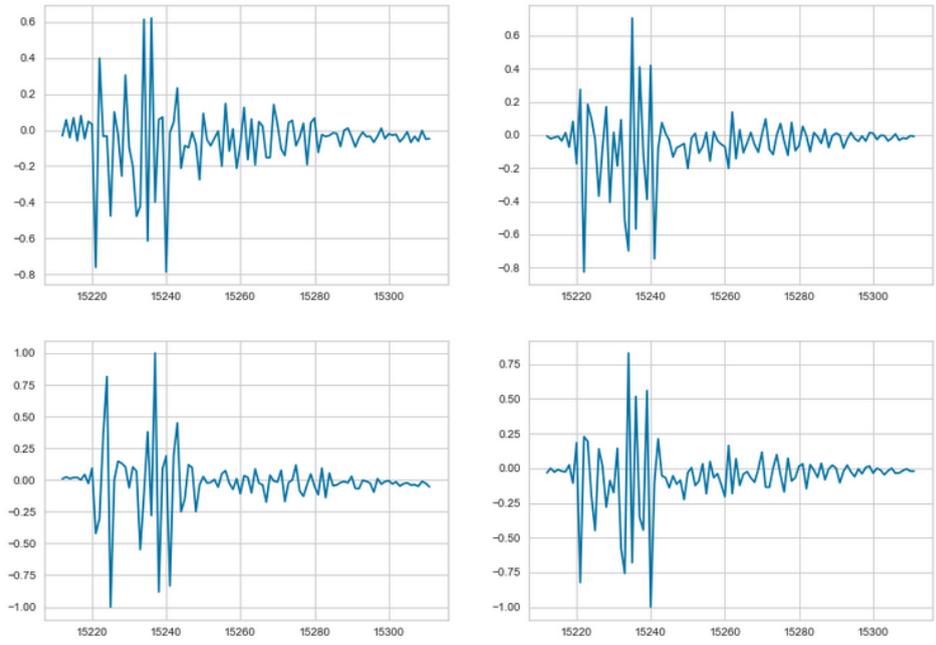


Figura 21: *Muestra espectro acústico, Plástico 7*

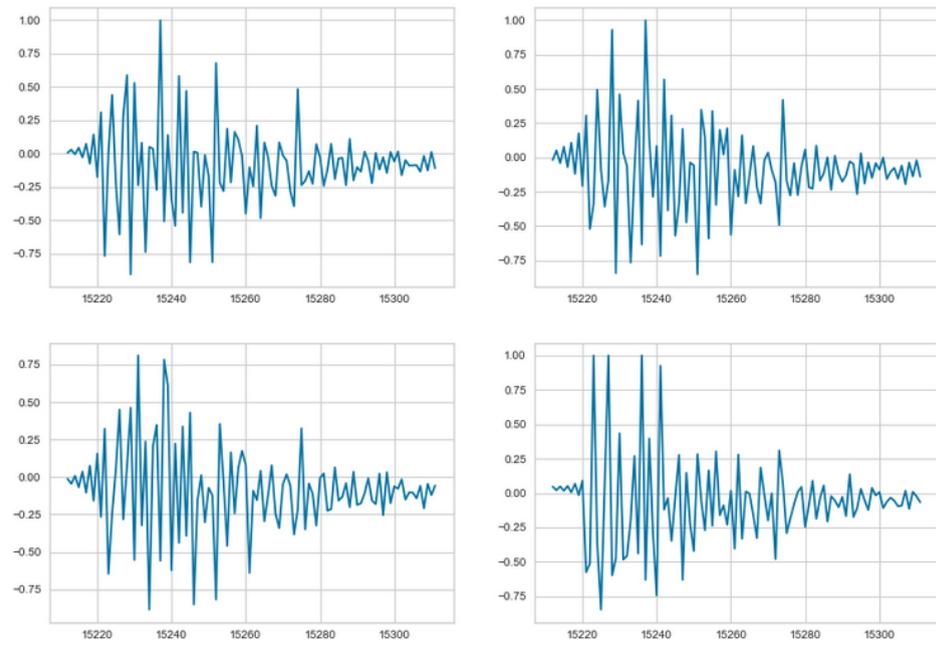


Figura 22: *Muestra espectro acústico, Plástico 8*

Al observar estas muestras, se podría concluir que el espectro acústico es irrelevante de cara a la clasificación. Pero observando con mayor detalle, la Figura 20 conteniente de muestras correspondientes al plástico 7 muestra claras características diferenciadoras con respecto al resto de plásticos. Más adelante se mostrará como se ha utilizado esta información para el clasificado.

3.2.3. Reducción de dimensionalidad - Espectro óptico

El rendimiento de los modelos de machine learning más comunes suelen verse afectados cuando son expuestos a un número grande de variables. Además de los problemas que puedan tener las relaciones entre dichas variables, como las correlaciones en el caso de regresiones lineales, los problemas causados al intentar tratar dicho número de variables han sido tantos que se ha llegado a denominar dicha situación como la *Maldición de la Dimensionalidad* (También conocido como *Curse of Dimensionality*), expresión forjada por primera vez por Richard E. Bellman al intentar tratar los problemas surgidos en programación dinámica.

Debido a esto, es deseable reducir las variables de entrada hasta un valor que sea adecuado para el modelo que vaya a utilizarse. Además, esto trae otros beneficios. Por ejemplo, al reducir la dimensión de los datos, es decir, aproximar una representación de los datos en un espacio de dimensión menor, normalmente se mantiene la información más relevante[2].

Toda la teoría de la reducción de dimensionalidad tiene por objetivo desarrollar herramientas que permitan reducir el número de variables conservando la mayor cantidad de información posible.

Para nuestro conjunto de datos de espectros ópticos, se han probado distintas técnicas de reducción de dimensionalidad. De estas estudiaremos como se distribuyen nuestros datos después del escalado y posteriormente un PCA. El método de escalado utilizado es el estándar es decir:

$$f(x_i) = \frac{x_i - \min(x)}{\max(x) - \min(x)} \quad (2)$$

Para cada x_i en el conjunto de variables.

La proyección del conjunto de datos sobre el plano determinado por la primera y segunda componente principal es el siguiente:

De manera independiente, se ha tomado la decisión de además de estandarizar los datos, realizar un análisis de componentes principales como parte del flujo de preprocesado de datos antes de llegar a la fase de modelado. Es necesario elegir el número de componentes utilizadas para la representación de los datos. Para ello, se hará uso del método del codo para la elección del número de componentes a incluir. Con este objetivo, a continuación se muestra un gráfico que representa el ratio de la varianza explicada frente a la n-ésima componente principal.

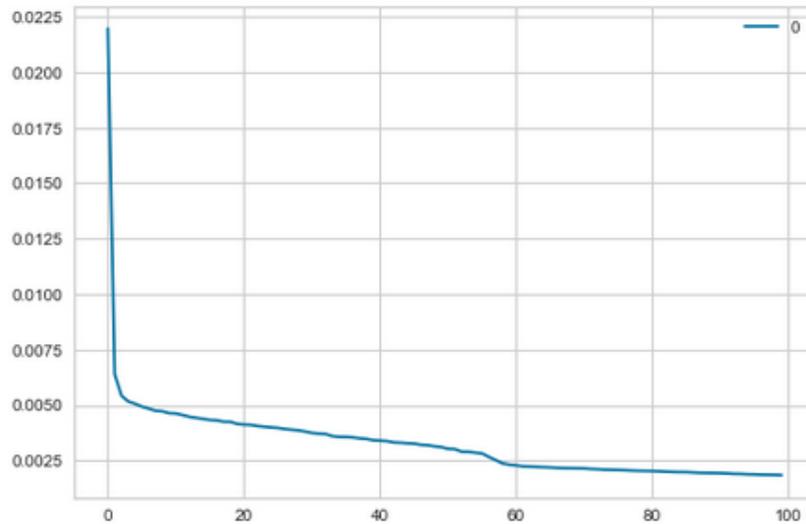


Figura 23: *Varianza explicada frente a la n-ésima componente principal*

El primer cambio de dirección se produce en la segunda componente principal. Sin embargo, no se ha considerado tomar las dos primeras componentes principales porque estas acumulan poco más del 2% de la varianza de los datos. También se puede observar que antes de la sexagésima componente principal también ocurre una disminución en el ratio de varianza acumulada en las componentes siguientes. Por tanto, se ha tomado la decisión de representar los datos mediante sus 60 primeras componentes principales.

Aunque en el entrenamiento de los modelos utilizemos estas 60 componentes, vamos a representar la proyección de los datos del espectro óptico sobre las dos primeras componentes principales.

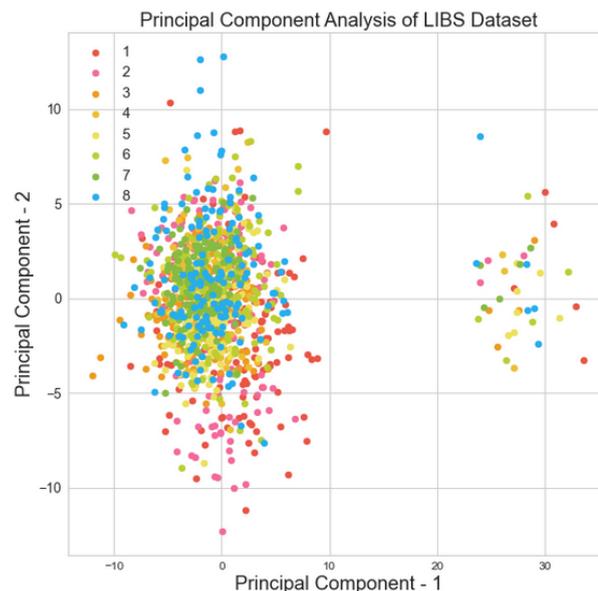


Figura 24: *Proyección de los datos en el plano Componente 1 - Componente 2 (ACP)*

Sobre este gráfico se destacan principalmente los siguientes puntos. Por un lado, vemos que las zonas no están claramente diferenciadas, es decir, observamos que existen varias muestras pertenecientes a categorías distintas en una misma zona del plano. Esto pudiera deberse a que dichas muestras están en una misma zona en el espacio vectorial representado por las componentes principales. Sin embargo, podría deberse también a que esta representación implica una reducción de dimensionalidad que puede no ser fiel a la realidad de los datos, siendo apenas sus proyecciones las que pertenecen a una misma zona del espacio.

Lo considerado anteriormente está de acuerdo con lo visto en la Figura 23, que muestra que el ratio de varianza acumulada por las dos primeras componentes principales se sitúa alrededor del 2%. Por tanto, no es de sorprender que estas dos componentes no basten para distinguir los tipos de plásticos. En los próximos apartados recogeremos la información sobre el entrenamiento y evaluación de los modelos en base a estas componentes.

3.3. Clasificación

El conjunto de datos fué entregado en un fichero con extensión ".csv". Este fichero contenía 15312 columnas. Una para la clase, 100 para el audio y 15211 para las intensidades ópticas. En el tratamiento previo se separaron estos datos según el tipo al que pertenecen.

En el conjunto de datos, la clase de cada muestra se encontraba enumerada del 1 al 8. Esta codificación es conocida como *Label Encoding*. Aunque es posible un entrenamiento con esta codificación, esto no es recomendable. Por ello, se ha realizado una transformación *One Hot Encoding* a la clase objetivo. Esto consiste en transformar la clase a la que pertenecen en un vector de dimensión igual al número de clases en las que solo hay un valor distinto de cero (e igual a uno) en la posición que represente a la clase a la que pertenece. Esto permite entrenar la red neuronal para que la salida sea un vector (o array) con las probabilidades de que la muestra pertenezca a cada clase.

Además, se ha dividido el conjunto de datos en dos subconjuntos, entrenamiento y test. El conjunto de entrenamiento está formado por 1089 muestras mientras que el conjunto de test está formado por 121 muestras. La diferencia entre los conjuntos utilizados para la clasificación basada en espectro óptico y la basada en espectro acústico es el conjunto de variables utilizado. Sin embargo, las muestras de los conjuntos son las mismas.

3.3.1. Clasificación a partir del espectro óptico

Conforme a lo considerado en el apartado 3.3 el conjunto base para esta clasificación contenía 1089 muestras. También consideramos en el apartado 3.2.3 que se elegirían las primeras 60 componentes principales para realizar la clasificación. Teniendo esto en consideración, las redes probadas se han intentado reducir al máximo, con el objetivo de evitar el sobreajuste (u overfitting).

La red final está formada por 2 capas de clasificación y una para la salida. Las capas de clasificación son densas, es decir, todas las neuronas están conectadas entre sí, entre capas. Además, se ha decidido que todas las activaciones son relu, salvo la capa de salida, que utiliza la activación softmax para la clasificación multiclase.

Model: "sequential_5"

Layer (type)	Output Shape	Param #
dense_20 (Dense)	(None, 400)	6084800
dense_21 (Dense)	(None, 100)	40100
dense_22 (Dense)	(None, 50)	5050
dense_23 (Dense)	(None, 8)	408

=====
Total params: 6,130,358
Trainable params: 6,130,358
Non-trainable params: 0
=====

Figura 25: *Arquitectura de la red entrenada para clasificación del espectro óptico*

Para verificar que la red ha sido correctamente entrenada, podemos observar la relación entre las funciones de pérdida en los conjuntos de entrenamiento y test durante el entrenamiento. En esta vemos que ambas están muy juntas y bajan progresivamente, lo que no da evidencias de sobreajuste ni una alta diferencia entre los conjuntos de entrenamiento y test. Además, vemos que la bajada ha sido progresiva, por lo que podría indicar una correcta selección del *learning rate*.

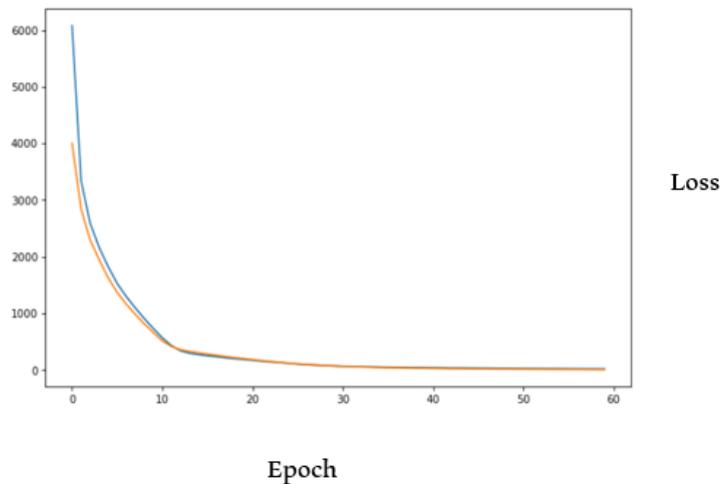


Figura 26: *Función de pérdida en los conjuntos de entrenamiento y validación en el conjunto obtenido por LIBS*

Además, vemos que las precisiones entre ambos conjuntos son similares, sin existir una diferencia significativa entre las precisiones de cada conjunto. Esto, una vez más indica un correcto funcionamiento y que se ha evitado el sobreajuste eficazmente.

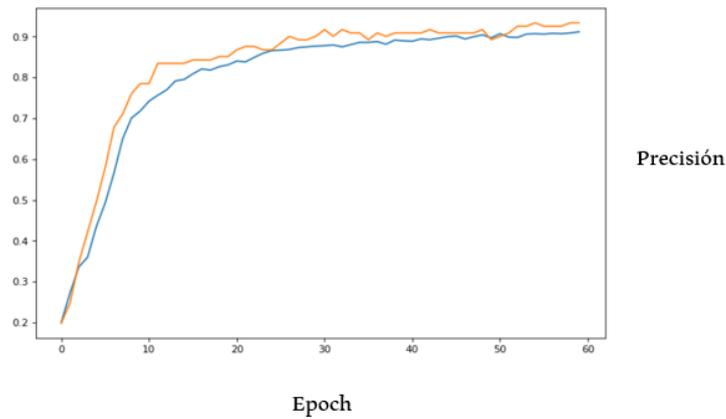


Figura 27: Precisión en los conjuntos de entrenamiento y validación de la red en el conjunto obtenido por LIBS.

Por último, en la matriz de confusión se observa algo destacable sobre el modelo y los datos. La matriz nos muestra como se distribuye la precisión del modelo entre las clases. Por lo que podemos observar que todas las clases se predicen con un porcentaje alto salvo las clases 4 y 7, que se confunden mutuamente. Esto puede deberse a que los plásticos son parecidos u a otros motivos desconocidos. Esta característica volverá a ser considerada junto a otros datos en el apartado de conclusiones.

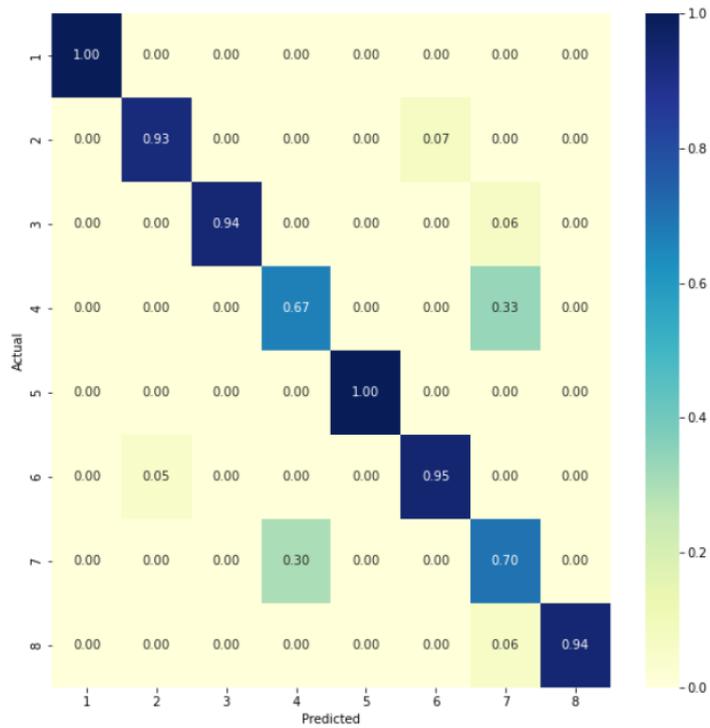


Figura 28: Matriz de confusión sobre el conjunto de test en el conjunto obtenido por LIBS

3.3.2. Clasificación a partir del espectro acústico

El espectro acústico tiene una complejidad añadida con respecto a otro tipo de variables utilizadas en el entrenamiento de modelos. Esto es, la translación en la muestra. Si representamos cada muestra de sonido mediante la notación t_1, \dots, t_{100} , siendo t_i la intensidad captada por el micrófono en el momento i , de manera óptima y en condiciones favorables, el vector (t_1, \dots, t_{100}) debería ser el mismo en todo momento. Sin embargo, podría ocurrir que no fuera así y que la muestra tuviese un desfase implicando que la muestra no fuera la misma. Esto podría tener consecuencias en métodos como el PCA, del cual no tiene sentido hacer uso si las variables no son las mismas.

Para prevenir este tipo de errores y con el objetivo de realizar una clasificación basada en las características implícitas de cada espectro acústico, se ha optado por utilizar una arquitectura en los modelos que hagan uso de redes convolucionales, en este caso, de una dimensión para que sean apropiadas para datos con estas dimensiones.

Conforme a lo mencionado anteriormente, las redes convolucionales están formadas por un número de filtros que se desplazan por la imagen con el objetivo de identificar las formas buscadas. El desplazamiento ayuda a identificar el objeto final buscado sin importar la zona de la imagen en la que esté posicionado. Para hacer que estas redes sean compatibles con nuestro espectro, el kernel, o filtro, debe tener una fila y n columnas, siendo n el número de momentos consecutivos para los cuales queremos comparar los valores de la muestra.

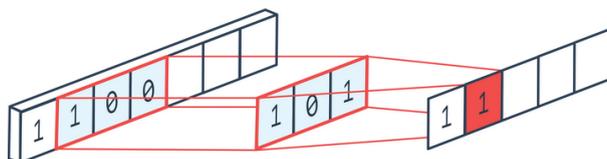


Figura 29: *Funcionamiento de un filtro lineal*

El diseño de la arquitectura de la red ha tenido como objetivo mantener la sencillez. Dado que el objetivo es descubrir si existe o no información útil para la clasificación en la intensidad sonora, la elección de una red más sencilla reduce la posibilidad de overfitting. Aunque esto implique que la red tenga una menor capacidad de aprendizaje, esto no tiene importancia en el contexto actual, teniendo en cuenta que el objetivo no es la construcción de un modelo infalible que vaya a ser desplegado para recibir peticiones.

La red diseñada se divide conceptualmente en 3 partes. Estas partes son: (1) 2 capas convolucionales y de pooling, (2) capa de Flatten (aplanado) y red densa y (3) capa softmax para la clasificación.

En primer lugar, las dos capas de pooling están compuestas por 3 filtros con forma 4×1 cada una. Estos filtros son los que se adaptan al input para extraer los comportamientos que puedan influir más en la clasificación. Seguido de cada capa convolucional está una capa de pooling. Esta secuencia de capas convolucional - pooling ha ganado protagonismo en la última década debido a su uso rendimiento en la clasificación

de imágenes y aparición entre las redes más conocidas como son la VGG16 y la VGG19.

Posteriormente, la capa Flatten es necesaria para transformar la forma de la salida de la última capa de pooling, formada por 3 vectores de dimensión 24, a un único vector de tamaño 24x3. Esto es necesario para que se pueda aplicar una última capa densa antes de la final, que devolviera un vector de dimensión 8 (uno por cada tipo de plástico), con las probabilidades de que la muestra tratada sea de cada una de las clases conocidas.

```

Model: "sequential_1"

```

Layer (type)	Output Shape	Param #
conv1d_2 (Conv1D)	(None, 100, 3)	15
max_pooling1d_2 (MaxPooling1D)	(None, 50, 3)	0
conv1d_3 (Conv1D)	(None, 48, 3)	30
max_pooling1d_3 (MaxPooling1D)	(None, 24, 3)	0
flatten_1 (Flatten)	(None, 72)	0
dense_2 (Dense)	(None, 12)	876
dense_3 (Dense)	(None, 8)	104

```

=====
Total params: 1,025
Trainable params: 1,025
Non-trainable params: 0
=====

```

Figura 30: *Arquitectura de la red entrenada para clasificación del espectro acústico*

Por otro lado, se puede observar el correcto entrenamiento de esta red mediante la figura mostrada a continuación, que representa el valor de la función de coste (o pérdida), para el conjunto de entrenamiento (en azul) y para el conjunto de validación (en naranja). Las curvas representadas tienen una apariencia bastante normal, siendo destacable la separación entre ambas. Esta separación es conocida como la varianza del modelo y representa un pequeño sobreajuste de este.

En condiciones óptimas, la función de pérdida en el conjunto de entrenamiento y el conjunto de validación debería ser igual. Sin embargo, esta variación se considera normal, dado que la igualdad pudiera indicar que el conjunto de entrenamiento y validación comparten al menos una parte de la muestra, lo que no sería óptimo para la evaluación del modelo.

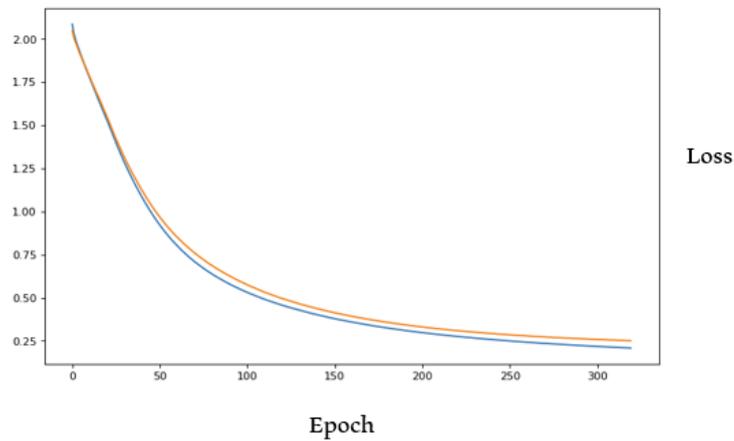


Figura 31: *Función de pérdida en los conjuntos de entrenamiento y validación de la red.*

Además, se muestra a continuación la precisión del modelo en los conjuntos dados. Se puede destacar que ambas precisiones son similares. Esto se debe en parte a la simplicidad de la red, lo que evita que se aprenda del ruido del conjunto de entrenamiento y fuerza a centrar su aprendizaje en comportamientos intrínsecos y más importantes de cada tipo de plástico. Por estos motivos, no hay gran diferencia entre las precisiones de los conjuntos de entrenamiento y test.

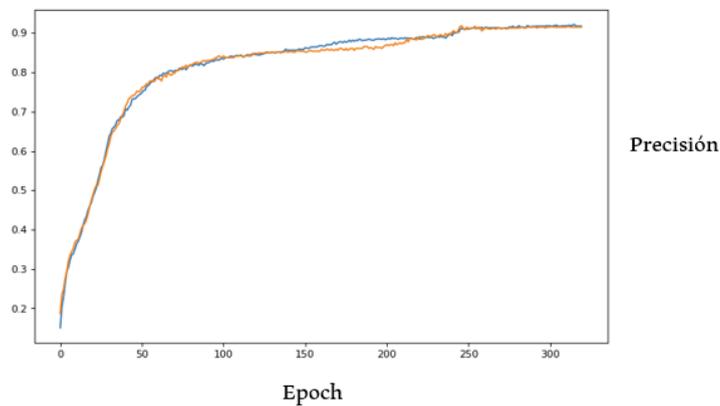


Figura 32: *Precisión en los conjuntos de entrenamiento y validación de la red en el conjunto acústico.*

Además, vemos mediante el gráfico que relaciona la precisión del modelo en los conjuntos de entrenamiento y test a medida que avanza el entrenamiento que son muy similares, lo que reafirma la idea de que no hay sobreajuste en este modelo.

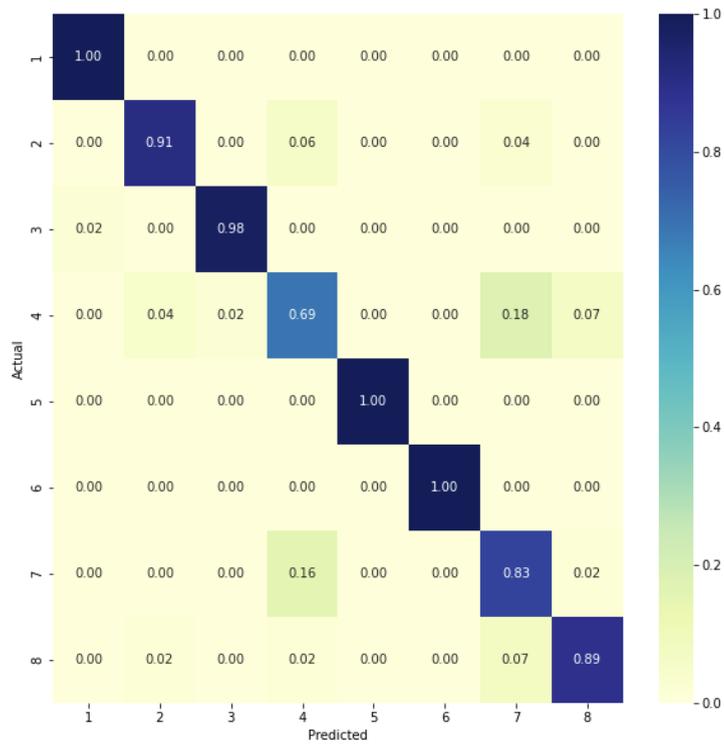


Figura 33: *Matriz de confusión sobre el conjunto de test en el conjunto acústico*

Por otro lado, gracias a la matriz de confusión podemos obtener una cierta relación entre las predicciones de los modelos. Al observar con detalle estas relaciones, vemos que las clases 4 y 7 tienen una probabilidad mayor de confusión entre ambos. De todas las clases, la 4 y la séptima son las que obtienen una menor precisión de clasificación.

4. Conclusiones

En primer lugar se ha logrado obtener un modelo de clasificación para cada conjunto de variables. Cabe destacar que ambos conjuntos obtienen una precisión superior al 90 %. Esto abre dos posibles líneas de investigación. La primera de ellas consiste en mejorar ambos modelos individualmente. La otra, considerar otro tipo de algoritmos basados en ensembles o alguna arquitectura que permita combinar los datos procesados mediante el PCA y la información acústica, pudiendo esto mejorar la clasificación global. Además, esto estaría relacionado con conocer si la información obtenida mediante el sonido es igual a la información obtenida mediante la técnica LIBS o si los modelos se basan en propiedades distintas para realizar la clasificación.

Además, un posible desarrollo se encuentra en comprobar la precisión de otros modelos existentes, como pudieran ser Random Forests, SVMs, XGBoost o Catboost. Esta comparación remarcaría las diferencias de una red neuronal con otro tipo de algoritmos.

En segundo lugar se ha observado una cierta relación entre las clases 4 y 7. Aunque este trabajo se ha realizado sin conocimiento sobre qué tipo de plásticos se han analizado, el hecho de que tanto el modelo basado en el espectro obtenido mediante LIBS como el modelo basado en la intensidad sonora confundieran las clases 4 y 7 pudiera indicar que estos plásticos tienen propiedades relacionadas, pudiendo llegar a ser de un tipo parecido.

Por este motivo, una vez planteadas estas hipótesis, se obtuvo información de los tipos de los plásticos que se estaban tratando. El tratar los datos de manera agnóstica colaboró para que obtener una opinión más objetiva. Los plásticos son los siguientes:

1. PVC espumado (no macizo), color blanco
2. Policarbonato transparente
3. Polipropileno blanco
4. Polietileno (PET) transparente
5. Acetal (POM) negro
6. PLA negro (de impresoras 3D)
7. Polietileno (PET) transparente
8. PLA negro (otro fabricante)

Con esta información, se puede llegar a entender la confusión del modelo a la hora de clasificar los plásticos 4 y 7. Que se confundan entre tipos muestra que los plásticos pueden tener similitudes al ser del mismo tipo. Para entender cuáles son las similitudes, se podría intentar seleccionar un conjunto de lambdas con la técnica “Forward Feature Selection” que permitiría identificar los conjuntos de longitudes de onda importantes para la clasificación. Sin embargo, esto solo serviría para unas primeras iteraciones dado que al tratarse de 15000 variables, la complejidad aumenta drásticamente.

Por otro lado, surge la duda de porqué el modelo no confunde los plásticos 6 y 8. En este caso sería apropiado identificar la composición de cada material, así como los posibles aditivos que pudieran ser incluidos por cada fabricante.

Durante el transcurso de este trabajo, se ha planteado que la información acústica pudiera estar condicionada a otros factores externos al tipo de plástico, como pudiera ser la forma y el grosor del plástico. Esto genera otra posibilidad de investigación, al procesar dichos materiales y repetir el análisis con plásticos con misma forma y grosor.

Por otro lado, y de cara a posibles avances más allá de este TFG, es importante considerar el color y el tipo de plástico. Esto se debe a que el tipo de plástico y el color puede afectar a la cantidad de luz absorbida por el material, lo que podría variar la intensidad del espectro recibido.

5. Bibliografía

Referencias

- [1] Osorio, Lesther Moreira, Luís V. Ponce Cabrera, and Eduardo de Posada. *Descubriendo la materia mediante luz láser: La espectroscopia de plasma inducido por láser (LIBS) como método para determinar la composición elemental de la materia*. Editorial Académica Española, 2012.
- [2] Wahid, Rahmat Abdul, Asiah Yahaya, and Yusof Munajat. *Detection of calcium in rice using laser-induced breakdown spectroscopy (LIBS)*. Malaysian Journal of Fundamental and Applied Sciences 10.2 (2014).
- [3] Pérez, A., et al. *Differentiation and Identification of Polymers using Laser-Induced Breakdown Spectroscopy (LIBS) in Combination with Multivariate Analysis Techniques*.
- [4] Murphy, Kevin P. *Machine learning: a probabilistic perspective*. MIT press, 2012.
- [5] Legnaioli, S., et al. *Industrial applications of laser-induced breakdown spectroscopy: a review*. Analytical Methods 12.8 (2020): 1014-1029.
- [6] Nair, Vinod, and Geoffrey E. Hinton. RECTIFIED LINEAR UNITS IMPROVE RESTRICTED BOLTZMANN MACHINES. Icm1. 2010.
- [7] Burgio, Lucia, et al. PIGMENT IDENTIFICATION IN PAINTINGS EMPLOYING LASER INDUCED BREAKDOWN SPECTROSCOPY AND RAMAN MICROSCOPY. Spectrochimica Acta Part B: Atomic Spectroscopy 56.6 (2001): 905-913.
- [8] Hahn, David W., and Nicoló Omenetto. *Laser-induced breakdown spectroscopy (LIBS), part II: review of instrumental and methodological approaches to material analysis and applications to different fields*. Applied spectroscopy 66.4 (2012): 347-419.
- [9] Singh, Vivek K., and Awadhesh K. Rai. *Potential of laser-induced breakdown spectroscopy for the rapid identification of carious teeth*. Lasers in medical science 26.3 (2011): 307-315.
- [10] Pavel, P., et al. *On the utilization of principal component analysis in laser-induced breakdown spectroscopy data analysis, a review*. Spectrochim. Acta Part B At. Spectrosc 148 (2018): 65-82.