

Aprendizaje automático aplicado a un modelo simplificado de detección de gases en imágenes hiperespectrales

Facundo F. Criscuolo: facundocriscuolo@gmail.com

Facultad de Ciencias Exactas y Naturales

Julio de 2025

Resumen

En esta monografía se desarrolló un modelo simplificado para la detección de gases mediante redes neuronales entrenadas con imágenes hiperespectrales sintéticas. Las imágenes fueron generadas a partir de espectros ficticios, perturbados mediante ruido aleatorio. Se diseñó una red neuronal de clasificación multiclase básica y se logró entrenarla exitosamente, alcanzando alta exactitud en la clasificación de gases conocidos. Se exploró la capacidad de generalización del modelo frente a variantes espectrales no vistas, evidenciando su sensibilidad a cambios sistemáticos no representados en los datos. Además, se ilustró el sobreajuste como consecuencia del aumento de la complejidad del modelo y la reducción de datos de entrenamiento. El trabajo permitió analizar de forma didáctica el comportamiento de redes neuronales simples aplicadas al análisis espectral y resaltar la relevancia del diseño del conjunto de datos para evitar sesgos y mejorar la robustez del modelo.

1. Introducción

Hoy en día, el aprendizaje automático constituye una herramienta fundamental para abordar problemáticas complejas en diversos campos, gracias a su capacidad para identificar patrones y realizar predicciones a partir de grandes volúmenes de datos.

En el ámbito de la teledetección, la disponibilidad creciente de datos adquiridos por sensores remotos, ha transformado la forma de monitorear y estudiar el ambiente. Este campo genera datos masivos y de alta dimensionalidad, como imágenes hiperespectrales, lo que habilita la aplicación de redes neuronales profundas como herramienta para extraer información relevante, reconocer patrones complejos y realizar predicciones precisas. Entre las aplicaciones más destacadas de la teledetección se encuentran el análisis del estado de suelos y cultivos, el monitoreo de la salud de la vegetación y el estudio de la evolución temporal de glaciares, entre muchas otras.

Un caso particular, que ha despertado interés creciente en la industria petrolera, es la detección y caracterización de emisiones gaseosas. Este desafío resulta especialmente relevante en operaciones de *oil & gas*, donde identificar y cuantificar emisiones puede contribuir a prevenir riesgos ambientales, mejorar la seguridad ocupacional y asegurar el cumplimiento de lineamientos normativos.

Dependiendo de su composición química, los gases pueden exhibir diferentes propiedades ópticas. Típicamente se los caracteriza a través de su espectro de transmisión (o de absorción), conocido como firma espectral, el cual es una medida de su transmitancia (o absorbancia) en función de la longitud de onda y es único para cada tipo de compuesto (ver figura 1) [1].

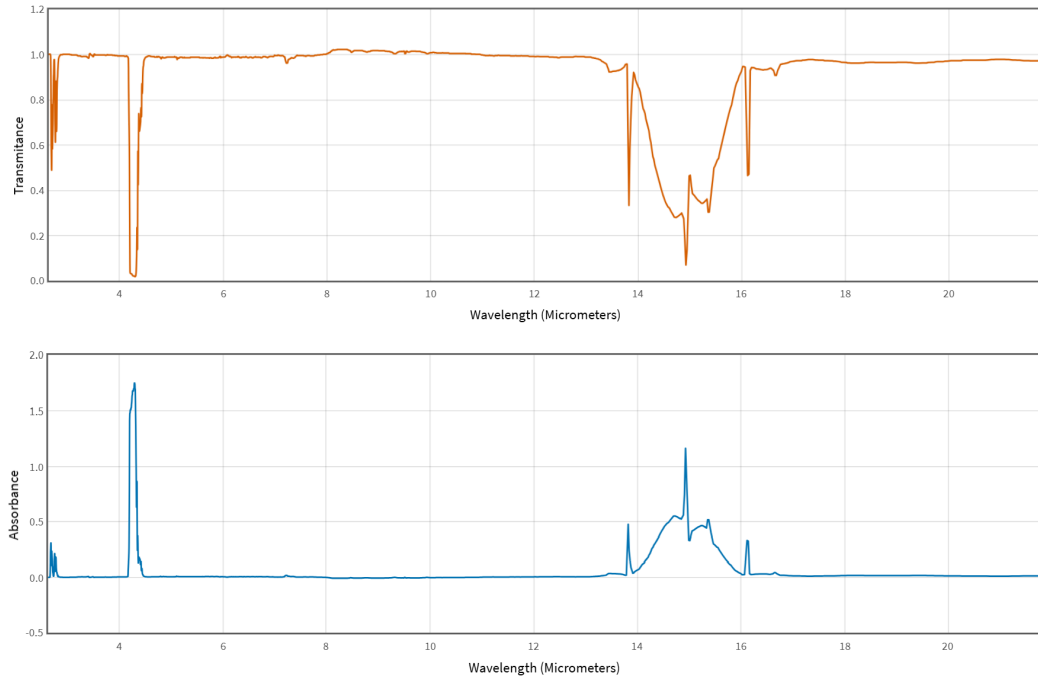


Fig. 1: espectros de transmisión (en naranja) y de absorción (en azul) para el CO_2 .

Los satélites empleados en teledetección óptica permiten captar las firmas espectrales de los gases registrando la reflectancia de los mismos en función de la longitud de onda, operando como espectrómetros. Cuando este registro se realiza en múltiples bandas espectrales contiguas se dice que la imagen resultante es hiperspectral.

Las imágenes (o cubos) hiperspectrales son arreglos tridimensionales donde cada píxel, además de tener una posición en el plano x-y de la imagen, tiene un espectro de reflexión asociado, representado por el valor de brillo para cada banda ubicada en el plano z (ver figura 2) [2].

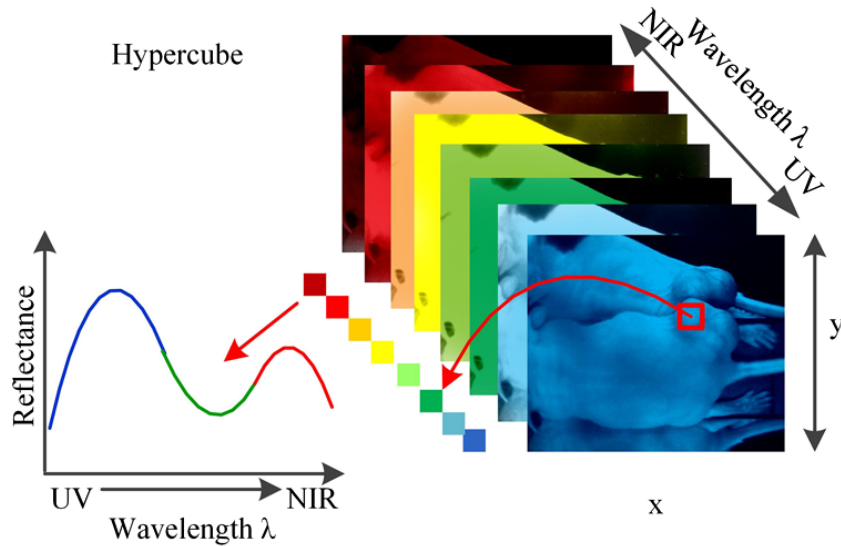


Fig. 2: cubo hiperspectral con firma espectral para un dado píxel.

En este contexto, las imágenes hiperspectrales permiten obtener información espectral detallada para cada píxel de una dada escena, facilitando la discriminación entre distintas firmas espectrales asociadas a diferentes gases. Combinadas con técnicas de aprendizaje automático, y en particular con redes neuronales profundas, se posibilita el desarrollo de modelos capaces de aprender a distinguir automáticamente entre diferentes compuestos.

La presente monografía aborda el diseño y la implementación de un modelo simplificado de detección de gases basado en el entrenamiento de una red neuronal utilizando imágenes hiperespectrales sintéticas. En particular, se busca discriminar entre tres firmas espectrales representativas, sentando las bases para futuras aplicaciones en escenarios reales y contribuyendo al desarrollo de herramientas automáticas para monitoreo ambiental.

2. Desarrollo

2.1. Datos

Se generaron 3 firmas espectrales “modelo” con una resolución de 20 bandas entre 200 nm y 900 nm a partir de distribuciones gaussianas centradas alrededor de las longitudes de onda correspondiente al azul, al verde y al rojo (ver figura 3).

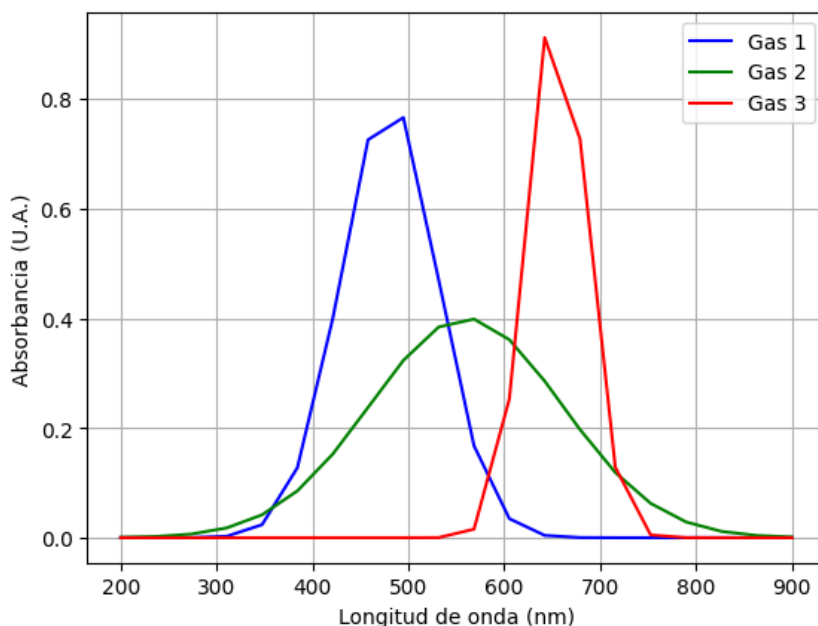


Fig. 3: firmas espectrales modelo para los gases 1 (en azul), 2 (en verde) y 3 (en rojo).

Para cada gas se construyeron 500 imágenes de 10 x 10 píxeles, y a cada píxel se le asignó el espectro del gas correspondiente modificado con cierto ruido, de modo que en total se construyeron 1500 imágenes hiperespectrales de 10 x 10 x 20. De esta manera, la absorbancia de un dado píxel para una longitud de onda ($A_{i,j}(\lambda)$) se vuelve resultado de perturbar con un ruido aleatorio $\epsilon_{i,j}$ a la absorbancia del gas n a esa misma longitud de onda ($A_n(\lambda)$):

$$A_{i,j}(\lambda) = A_n(\lambda)(1 + \epsilon_{i,j})$$

Los valores de ruido $\epsilon_{i,j}$ fueron generados para cada píxel usando la función *random.normal* de *numpy* y un valor de 0,4 para el desvío estándar.

Para visualizar algunas de las imágenes generadas en RGB se seleccionaron las longitudes de onda correspondientes al azul, verde y rojo y se asignaron a los canales correspondientes (ver figura 4).

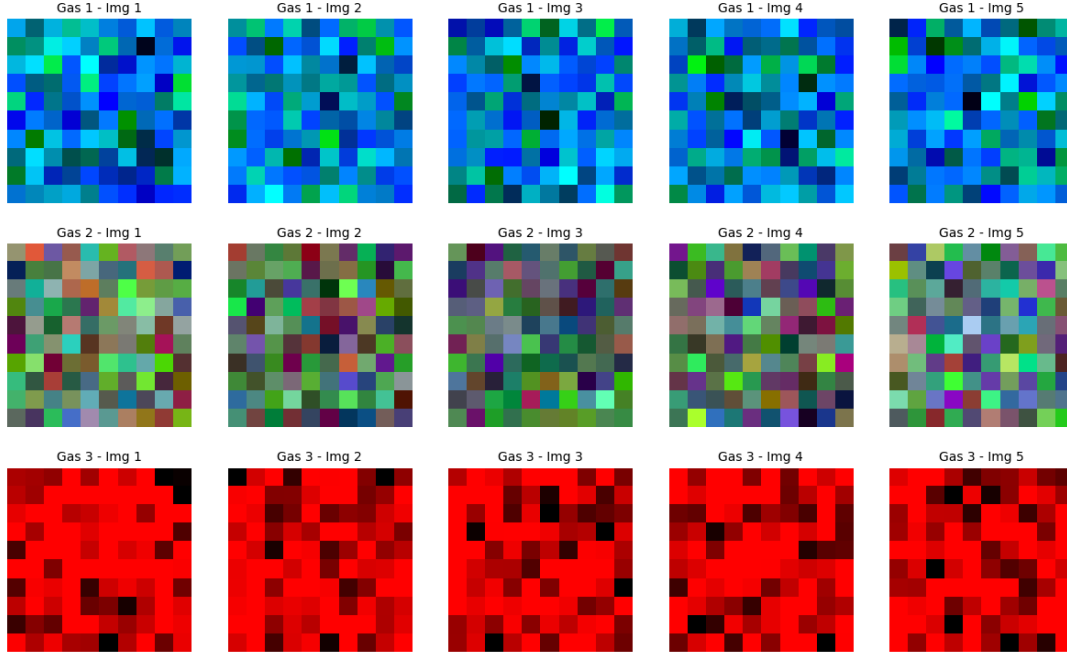


Fig. 4: primeras 5 imágenes simuladas para cada gas.

En la figura 4 se observa cómo en cada gas predomina el color correspondiente a su máximo de absorción, ya que en promedio los píxeles presentan valores de brillo más altos en dichas longitudes de onda.

Como el problema a resolver implica identificar al gas a partir de su firma espectral, ya que esta es su característica distintiva, y no a través de un patrón dado de píxeles, resulta necesario extraer un espectro de absorción promedio de cada imagen, para así entrenar a la red con datos de absorbancia en función de la longitud de onda. Haciendo esto para todas las imágenes, se puede observar que los espectros para un mismo gas resultan similares, con mayores diferencias para las longitudes de onda donde presentan mayores valores de absorbancia (ver figura 5).

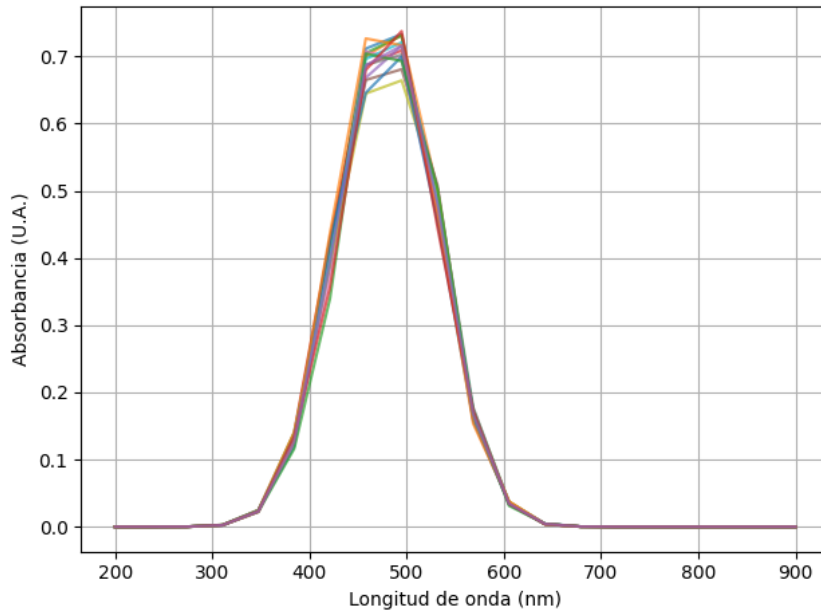


Fig. 5: primeros 15 espectros de absorción correspondientes al gas 1 obtenidos a partir de promediar los valores de brillo de los píxeles de las imágenes hiperespectrales generadas.

2.2. Arquitectura

A los datos obtenidos a partir de las imágenes se les asignó su clase correspondiente según el número de gas empleando *to_categorical* de *keras*, la cual genera una matriz binaria de $n \times n$ donde n es el número de clases. Con el 70 % del conjunto de datos se armó el conjunto de entrenamiento, correspondiendo el 30 % restante al conjunto de prueba.

En lo que respecta a la arquitectura de la red, la forma del *input* necesariamente queda determinada por la dimensión de los datos, los cuales son vectores de 20 valores (correspondientes a las absorciones de cada una de las 20 bandas). Al mismo tiempo, la forma del *output* se corresponde con el número de clases empleado para la clasificación, siendo la función de activación de dicha capa una *Softmax* a modo de obtener las probabilidades asociadas a que el gas pertenezca a cada una de las 3 especies. Para las capas ocultas, una posibilidad es usar funciones de activación *ReLU* reduciendo la dimensionalidad del *input*.

Inicialmente se propuso una arquitectura con una sola capa oculta de 5 unidades, usando al momento de compilar el modelo un optimizador *Adam* con un *learning rate* de 0,01 y *categorical cross-entropy* como función de pérdida, ya que el problema se trata de una clasificación multiclase (ver figura 6).

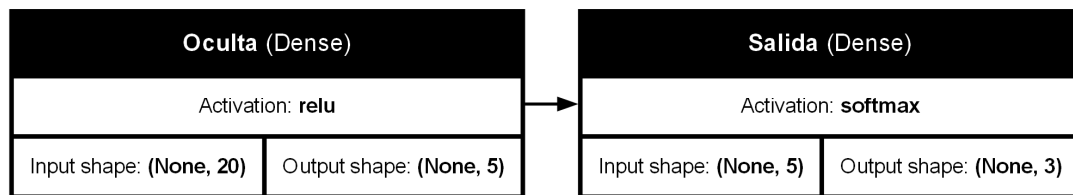


Fig. 6: modelo inicial propuesto para la arquitectura de la red, con una única capa oculta de 5 unidades.

En la figura 6 el valor “None” indica que la dimensión correspondiente al tamaño del *batch* no está fijada durante la definición del modelo, sino que es determinada durante el entrenamiento, permitiendo procesar cualquier cantidad de muestras.

3. Resultados

Habiendo entrenado al modelo en 6 épocas, se realizaron gráficos para la función de pérdida y para la exactitud de los conjuntos de entrenamiento y de prueba en función del número de época (ver figura 7).

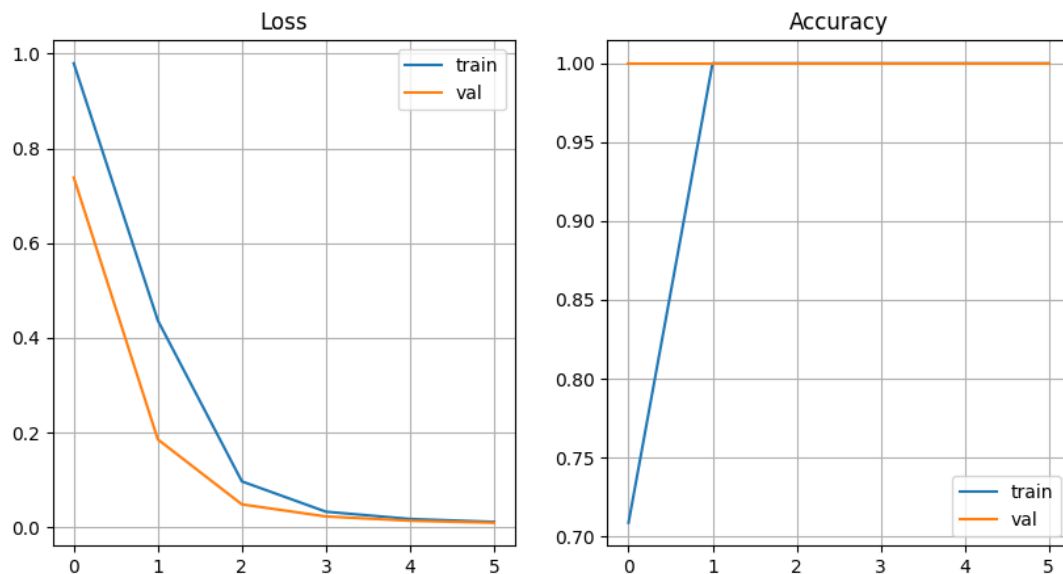


Fig. 7: a la izquierda función de pérdida (*Loss*) y a la derecha exactitud (*Accuracy*) en función del número de época, para el conjunto de entrenamiento (*train* en azul) y de prueba (*val* en naranja).

En la figura 7 se observa que el modelo rápidamente maximiza la exactitud en ambos conjuntos de datos, mientras que la función de pérdida tiende progresivamente a 0 hacia la última época. Incluso, que la función de exactitud se maximice tan rápido, probablemente sea resultado de la falta de complejidad en los datos generados. Todo esto parece indicar que la arquitectura propuesta resulta adecuada para tratar el problema simplificado de clasificación de gases.

3.1. Predicciones

A modo de probar el modelo entrenado se generó un nuevo gas perturbando aleatoriamente el espectro del gas 1 en cada banda (ver figura 8).

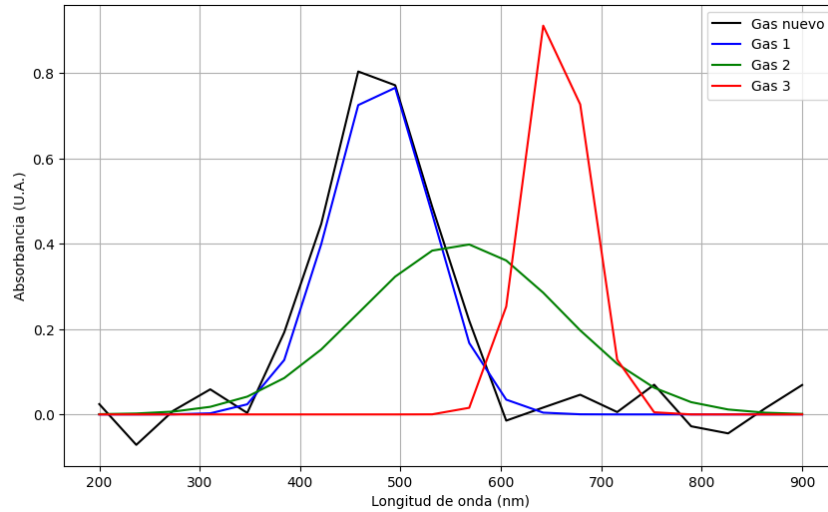


Fig. 8: firmas espectrales modelo para los gases 1 (en azul), 2 (en verde), 3 (en rojo) y para el gas nuevo (en negro).

En la figura 8 se observa que el gas nuevo presenta una alta similitud con el gas 1, pero que aún así toma valores de absorbancia distintos de cero por fuera de la campana central.

Usando la función *predict* sobre el espectro del gas nuevo, el modelo arroja una lista con la probabilidad que tiene el gas de pertenecer a cada una de las clases, en este caso:

$$[9,9478203 \times 10^{-1}; 4,8953141 \times 10^{-3}; 3,2262673 \times 10^{-4}]$$

Así el modelo predice exitosamente que el gas nuevo es el gas 1, con una probabilidad mayor a 0,99.

Sin embargo el modelo no contempla, por ejemplo, la posibilidad de que haya un *offset* en los espectros de absorción, ya que el entrenamiento se llevó a cabo con gaussianas con el “piso” en cero. Si el espectro estuviera solapado con una señal de base (como podría ser un ruido de fondo), habría que remover esta antes de intentar predecir a qué gas pertenece, ya que de lo contrario el modelo no podrá predecir su clase (ver figura 9).

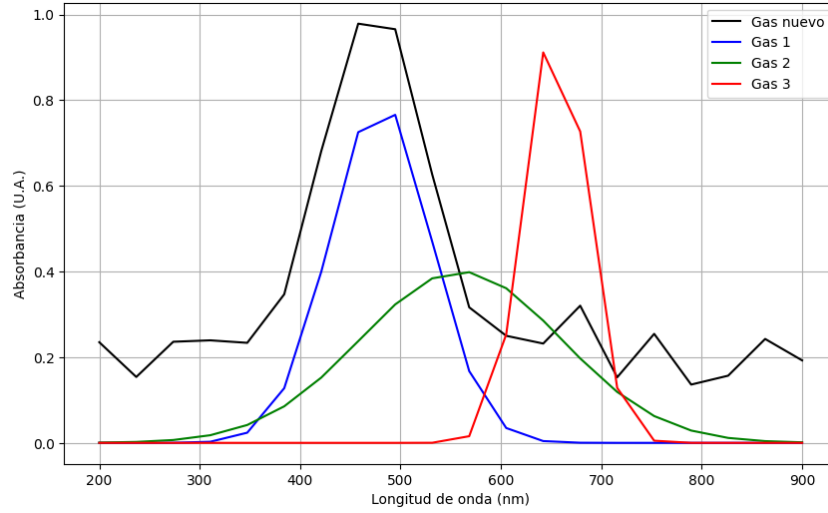


Fig. 9: firmas espectrales modelo para los gases 1 (en azul), 2 (en verde), 3 (en rojo) y para el gas nuevo con un *offset* (en negro).

En la figura 9 se observa que el máximo de absorción del gas nuevo coincide con el del gas 1, a pesar de estar corrido uniformemente en el eje de absorbancia.

En este caso, el modelo arroja la siguiente lista con la probabilidad que tiene el gas de pertenecer a cada una de las clases:

$$[6,7373317 \times 10^{-6}; 9,9998653 \times 10^{-1}; 6,7017208 \times 10^{-6}]$$

Así se ve que el modelo se ve condicionado por el *offset* introducido en la señal, ya que predice que el gas nuevo se corresponde con el gas 2. Esto refuerza la idea de preparar adecuadamente los datos antes de introducirlos ciegamente al modelo esperando una predicción acertada, o incluso de entrenar al modelo con un conjunto de datos representativo de las situaciones que se podrían encontrar en nuevas mediciones.

3.2. Sobreajuste

Un caso particular que se suele dar cuando la arquitectura del modelo es más grande de lo necesaria (muchas capas o muchas unidades) es el sobreajuste (u *overfitting*). Bajo esta condición el modelo ajusta los pesos y los *biases* a modo de clasificar perfectamente los datos del conjunto de entrenamiento, “memorizándolos” y luego no pudiendo clasificar correctamente los datos del conjunto de prueba.

Para visualizar este fenómeno se modificó la arquitectura de la figura 6, incorporando más capas y unidades, así como también se redujo la cantidad de datos del conjunto de entrenamiento, pasando a estar conformado por el 0,01 % del total de los datos (ver figuras 10 y 11).

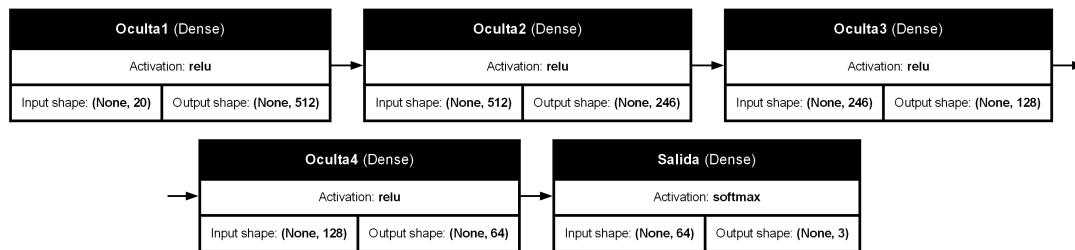


Fig. 10: modelo modificado propuesto para visualizar el sobreajuste.

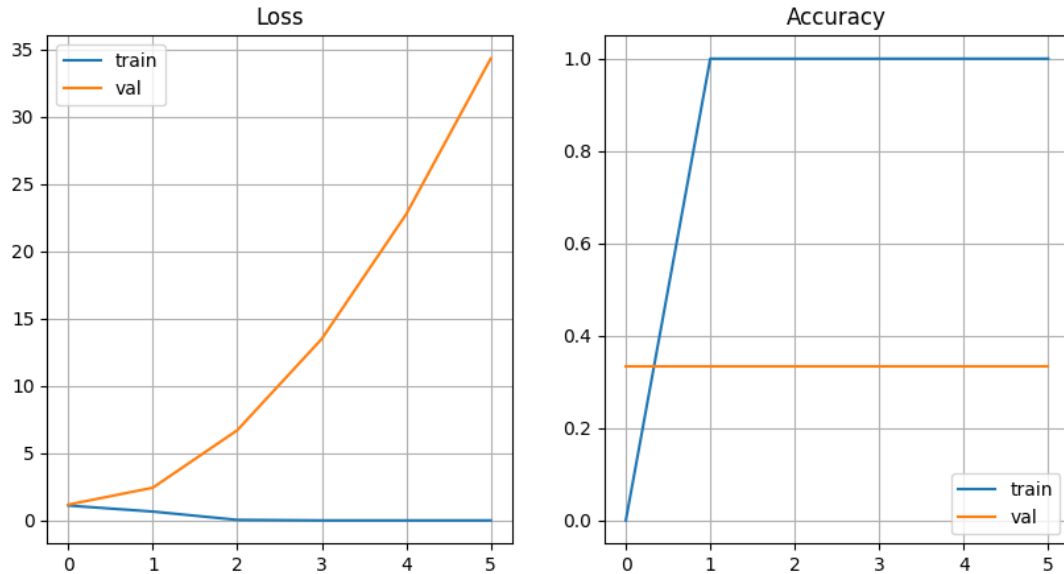


Fig. 11: a la izquierda función de pérdida (*Loss*) y a la derecha exactitud (*Accuracy*) en función del número de época, para el conjunto de entrenamiento (*train* en azul) y de prueba (*val* en naranja) modificados para visualizar el sobreajuste.

En la figura 11 se observa que la exactitud del conjunto de entrenamiento rápidamente se maximiza, mientras que en el conjunto de prueba se estanca en un valor menor a 0,4. La función de pérdida, si bien converge a 0 para el conjunto de entrenamiento, se dispara para el conjunto de prueba, como es de esperar para el caso de sobreajuste. Esto evidencia cómo una red puede memorizar, en lugar de generalizar, si no se cuida la selección de los hiperparámetros o la cantidad de datos.

4. Conclusiones

En este trabajo se implementó un modelo simplificado de detección de gases a partir del entrenamiento de una red neuronal utilizando imágenes hiperespectrales sintéticas. Las imágenes fueron generadas a partir de firmas espectrales de tres sustancias ficticias, con la adición de ruido aleatorio para simular variabilidad espectral. A partir de estas imágenes se construyó un conjunto de datos empleando los espectros promedio de cada imagen.

Se diseñó una red neuronal de clasificación multiclase de arquitectura simple: una capa de entrada de 20 unidades, una capa oculta de 5 neuronas con activación *ReLU*, y una capa de salida con 3 unidades y activación *Softmax*. Tras un entrenamiento breve (6 épocas), el modelo logró minimizar la función de pérdida y alcanzar una alta exactitud en la clasificación.

El modelo demostró capacidad de generalización al identificar correctamente una variante de una de las sustancias originales, con una probabilidad de clasificación superior al 0,99. No obstante, al modificar sistemáticamente dicho espectro mediante un *offset*, el modelo no fue capaz de reconocerlo, lo que evidenció su sensibilidad a transformaciones no representadas en los datos de entrenamiento.

Finalmente se ilustró el fenómeno de sobreajuste, aumentando la capacidad del modelo y reduciendo la cantidad de datos de entrenamiento. Como era de esperar, la red ajustó perfectamente los datos de entrenamiento (pérdida próxima a cero) pero falló en generalizar a los datos de prueba, mostrando una pérdida creciente y exactitud constante y limitada.

Referencias

- [1] Dow Chemical Company. Libro del Web de Química del NIST, SRD 69. 1964.
- [2] Sowmya Vishvanathan, Soman Kp, and Mahmoud Hassaballah. *Hyperspectral Image: Fundamentals and Advances*, pages 401–424. 01 2019.