

Uso de redes neuronal convolucionales 1D en espectrometrías para clasificación de géneros de microalgas

González, J.^{a,*}, Guzmán, J.L.^a, Ación, F.G.^b, Ciardi, M.^b, Moreno, J.C.^a

^aDep. de Informática, Universidad de Almería, CIESOL, ceiA3, 04120, España, {jgh043,joseluis.guzman,jcmoreno}@ual.es

^bDep. de Ingeniería Química, Universidad de Almería, CIESOL, ceiA3, 04120 Almería, España, {facien, mc722}@ual.es

To cite this article: González Hernández, J., Guzmán Sánchez, J.L., Ación Fernández, F.G., Ciardi, M., Moreno Ubeda, J.C. 2023. Use of convolutional 1D network in spectrometry for classification of microalgae genres. XLIV Jornadas de Automática, 399-404. <https://doi.org/10.17979/spudc.9788497498609.399>

Resumen

En este trabajo se presenta el desarrollo de una red neuronal para la clasificación de especies de microalgas basada en muestras espectrales. Los datos necesarios para realizar la clasificación son obtenidos mediante un espectrofotómetro, utilizado barridos espectrales como datos representativos para la caracterización de las muestras. El modelo desarrollado ha sido entrenado en base a estos datos y permite distinguir y clasificar entre 4 géneros diferentes de microalgas, *Spirulina*, *Chlorella*, *Synechococcus* y *Scenedesmus*. Los resultados demuestran una elevada tasa de acierto y precisión en el modelo frente a muestras puras, presentando un gran potencial para el desarrollo futuro de herramientas para la monitorización de cultivos a gran escala que permita contribuir al control de contaminación del cultivo.

Palabras clave: Aprendizaje computacional, Red Convolucional, Microalgas

Use of 1D convolutional network in spectrometry for classification of microalgae genres.

Abstract

This work presents the development of a neural network for the classification of microalgae species based on spectral samples. The data necessary to perform the classification are obtained by means of a spectrophotometer, using spectral scans as representative data for the characterization of the samples. The developed model has been trained based on these data and allows to distinguish and classify between 4 different genera of microalgae, *Spirulina*, *Chlorella*, *Synechococcus* and *Scenedesmus*. The results demonstrate a high hit rate and accuracy in the model against pure samples, presenting a great potential for the future development of tools for large-scale culture monitoring to contribute to the control of culture contamination.

Keywords: Machine Learning, Convolutional networks, Microalgae

1. Introducción

Las microalgas son microorganismos fotosintéticos que han despertado un gran interés en los últimos años debido a su potencial para abordar diversos desafíos ambientales y alimentarios (Ación et al., 2012). Dichos microorganismos unicelulares tienen la capacidad de reproducirse y crecer rápidamente en una amplia variedad de entornos (Tapie and Bernard, 1988), lo que los convierte en una valiosa fuente de soluciones sostenibles a problemas ambientales apremiantes en la actualidad como es el

tratamiento de agua residual (Nordio et al., 2022). Las microalgas ofrecen una solución prometedora en este campo. Estos organismos pueden eliminar contaminantes y nutrientes no deseados presentes en el agua residual a través de su capacidad de absorción y asimilación. Además, durante su crecimiento, las microalgas capturan dióxido de carbono (CO_2) de la atmósfera y lo convierten en biomasa, lo que ayuda a mitigar el cambio climático al reducir las emisiones de gases de efecto invernadero (Zhao and Su, 2014).

*Autor para correspondencia: jgh043@ual.es

Las microalgas también tienen un enorme potencial en el campo de la alimentación y la nutrición. Son ricas en proteínas, lípidos, vitaminas y minerales esenciales, lo que las convierte en una fuente valiosa de nutrientes (Spolaore et al., 2006). Por otro lado, algunas especies de microalgas contienen compuestos bioactivos con propiedades antioxidantes y (Kumar et al., 2022), lo que las hacen ideales para la producción de alimentos funcionales y suplementos nutricionales.

La producción de microalgas puede llevarse a cabo mediante diversos sistemas, incluyendo fotobiorreactores cerrados y cultivos exteriores (Guzmán et al., 2021). Los fotobiorreactores cerrados ofrecen un mayor control de las condiciones de cultivo, lo que puede resultar en una mayor productividad y calidad de las microalgas. Por otra parte, los cultivos exteriores, como estanques y lagunas, son más económicos y adecuados para la producción a gran escala (Banerjee and Ramaswamy, 2017; Villalobos et al., 2019). Sin embargo, uno de los desafíos en los cultivos exteriores de microalgas es la contaminación del cultivo. En un entorno abierto, existe el riesgo de que otros géneros de microalgas no deseadas puedan colonizar y competir con las especies de interés. Esto puede afectar la calidad y la productividad de los cultivos. Por tanto, es fundamental mantener un control y monitoreo del cultivo con el fin de evitarlo.

En la actualidad este proceso se lleva a cabo mediante análisis manual de las muestras bajo la inspección visual de un experto capaz de discernir las diferencias morfológicas entre géneros mediante un microscopio óptico (Roy et al., 2022), lo cual requiere de trabajadores altamente cualificados. Para facilitar este proceso, se ha estudiado el uso de FlowCam para la adquisición de imágenes de microalgas y redes neuronales para clasificarlas (Otálora Berenguel et al., 2021). Esto ha permitido replicar el proceso realizado en laboratorio, pero de manera más rápida y precisa, aunque con la desventaja de que el uso de FlowCam requiere una inversión elevada. Otro método alternativo es el análisis de la composición bioquímica de la biomasa, donde los pigmentos o perfiles de ácidos grasos permite realizar un control de la composición de la biomasa, pero no determina la presencia de otros géneros (Kozlova et al., 2017).

En este trabajo, se buscó desarrollar un enfoque alternativo y más económico que el FlowCam para la clasificación de géneros de microalgas. Para ello, se utilizó un espectrómetro óptico para obtener datos espectrales en el rango de 300 - 1000 nm. Este tipo de instrumento es comúnmente utilizado en laboratorios para el análisis de nutrientes en cultivos, por lo que su disponibilidad es alta en las plantas de cultivo de microalgas. Se propuso emplear una red neuronal basada en capas convolucionales 1D para procesar los datos espectrales y realizar la clasificación de los géneros de microalgas. Las capas convolucionales 1D son especialmente adecuadas para trabajar con datos secuenciales, como en este caso, donde cada muestra es un espectro.

La red neuronal fue entrenada utilizando un conjunto de datos previamente etiquetados, que incluía espectros de diferentes géneros de microalgas. Durante el entrenamiento, la red aprendió a identificar patrones y características específicas en los es-

pectros que permitían distinguir entre los diferentes géneros. Una vez entrenada, fue capaz de clasificar nuevas muestras de microalgas basándose en sus espectros. Este enfoque demostró ser prometedor en términos de precisión y eficiencia, al mismo tiempo que evitaba la adquisición de instrumentación adicional.

2. Materiales y métodos

En esta sección se describen las especies, así como el hardware utilizado para la realización de la espectrometría y el software seleccionado para la implementación de la red neuronal convolucional.

2.1. Especies

Para el entrenamiento de la red se emplearon cuatro especies disponibles en las instalaciones del centro IFAPA (La Cañada), una por cada género a clasificar, con morfologías diferentes: *Spirulina platensis*, *Chlorella Vulgaris*, *Synechococcus elongatus* y *Scenedesmus Almeriensis* cultivadas en laboratorio. Cada una de estas especies presenta características particulares y se utiliza en diversas aplicaciones debido a sus propiedades biológicas y químicas.

El género de *Spirulina* es ampliamente conocida por su alto contenido de proteínas, vitaminas y minerales. Esta microalga se ha utilizado tradicionalmente como suplemento alimenticio debido a su valor nutricional y beneficios para la salud (Spolaore et al., 2006). Sus propiedades antioxidantes, inmunomoduladoras y antimicrobianas, la convierte también en un objeto de interés en la industria alimentaria y farmacéutica (Kumar et al., 2022). *Chlorella* es reconocida por su alta capacidad de acumular lípidos y su contenido en ácidos grasos poliinsaturados (Kumaran et al., 2023). Además, posibles efectos beneficiosos en la salud, como la reducción de los niveles de colesterol y la mejora del sistema inmunológico están siendo investigados (Barghchi et al., 2023). *Synechococcus* es un género apreciada por su mecanismo eficaz para adaptarse a los cambios de salinidad e intensidad luminosa (Kim et al., 2018). *Scenedesmus* ha despertado interés en la producción de biodiesel y el tratamientos de aguas residuales con el fin de reducir su contenido en nitrógeno y fósforo (Nordio et al., 2022; Park et al., 2011).

2.2. Espectrofotómetro

En el presente estudio, se llevaron a cabo barridos espectrales utilizando un espectrofotómetro Genesys 10S UV-VIS (ver Figura 1) ubicado en el laboratorio del centro IFAPA (La Cañada), localizado cerca de la Universidad de Almería. Los barridos se realizaron en un rango de longitud de onda de 300 nm a 1000 nm. Este rango de longitud de onda fue seleccionado para abarcar tanto el espectro ultravioleta A como el visible, permitiendo así la caracterización de muestras en esta amplia gama de longitudes de onda. Los barridos espectrales se realizaron haciendo uso de cubetas de cuarzo para permitir el paso de la luz UVA y siguiendo un protocolo establecido, donde se registraron las lecturas de absorbancia con pasos de 5 nm. Esta metodología nos permitió obtener un perfil completo de absorción de las muestras en el rango de interés, brindando información detallada sobre las propiedades ópticas de los compuestos analizados.



Figura 1: Espectrofotómetro Genesis 10S UV-VIS

2.3. Python y librerías

Para la realización del trabajo se utilizó como lenguaje de programación *Python*, debido a su sintaxis simple y legible, así como a su amplia variedad de bibliotecas y herramientas especializadas en el procesamiento y análisis de datos. Dentro de las bibliotecas utilizadas destacan *Pandas*, que proporciona estructuras de datos y herramientas de análisis de datos de alto rendimiento. Es ampliamente utilizada para el procesamiento y manipulación de datos tabulares, como hojas de cálculo o bases de datos, facilitando la limpieza, filtrado, agregación y transformación de los datos. Además se hizo uso de *Keras*, que permite el desarrollo y entrenamiento de redes neuronales. Ofrece una interfaz sencilla y fácil de usar para construir modelos de aprendizaje profundo, diseñar arquitecturas de redes neuronales personalizadas, especificar capas y funciones de activación, y gestionar el proceso de entrenamiento y evaluación del modelo.

Para la evaluación del modelo se ha utilizado la técnica de validación cruzada *K-Fold* que permite evaluar y validar modelos de manera más rigurosa (Anguita et al., 2012). Consiste en dividir el conjunto de datos en K particiones. Luego, se entrena y evalúa el modelo K veces, utilizando una partición diferente como conjunto de prueba en cada iteración y el resto de las particiones como conjunto de entrenamiento. Al finalizar las K iteraciones, se promedian los resultados de evaluación para obtener una medida más robusta del rendimiento del modelo. El uso de validación cruzada ayuda a mitigar el sobreajuste y proporciona una evaluación más confiable del modelo en datos no tratados. Todo el proceso de entrenamiento y validación se ha realizado en un equipo con procesador *Ryzen 9 5900HX* y 32 Gb RAM DDR4.

2.4. Red neuronales convolucionales

Las redes neuronales convolucionales 1D (1D CNN) son una variante de las redes neuronales convolucionales diseñadas específicamente para el procesamiento de datos unidimensionales, como secuencias de texto (Kim, 2014), series temporales o señales de audio (Abdoli et al., 2019). Estas redes son muy efectivas en tareas de clasificación, detección de patrones y extracción de características en datos secuenciales (Ko et al., 2022).

El funcionamiento de las redes neuronales convolucionales 1D se basa en la aplicación de filtros convolucionales a lo largo de una dimensión de los datos de entrada, normalmente la dimensión temporal. Estos filtros se deslizan sobre la secuencia de entrada, calculando diferentes sumas ponderadas de los valores de entrada. A medida que el filtro se desplaza, se generan características locales que capturan patrones específicos en la secuencia, tal y como se muestra en la Figura 2

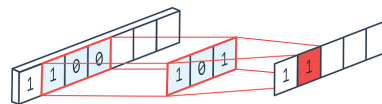


Figura 2: Funcionamiento de un filtro convolucional 1D

Una vez que se han extraído las características mediante las capas convolucionales, se utilizan capas de *pooling* para reducir la dimensionalidad y resumir la información relevante. El *pooling* en las redes 1D CNN generalmente se realiza mediante la operación de *max pooling*, que selecciona el valor máximo en cada ventana de *pooling*.

Las características se alimentan a través de capas completamente conectadas, conocidas como capas densas, para realizar la clasificación o tarea deseada. Estas capas pueden incluir funciones de activación para introducir no linealidad en el modelo. Finalmente, en el caso de problemas de clasificación multiclase, se incluye una función de activación *softmax* con el fin de obtener una distribución de probabilidades. Este proceso garantiza que los valores de salida estén en el rango de 0 a 1 y que la suma de todos los elementos sea igual a 1, lo que permite interpretar los resultados como probabilidades.

3. Resultados

En primer lugar, se realizaron diversos barridos espectrales en las distintas especies mencionadas en la sección anterior. Para la realización de dichos barridos se utilizó el espectrofotómetro Genesis 10S UV-VIS junto con cubetas de cuarzo para llevar a cabo los barridos espectrales en el rango de longitud de onda de 300 nm a 1000 nm. Este dispositivo es un instrumento altamente preciso y confiable, diseñado específicamente para la caracterización de muestras en dicho espectro.

Se empleó una cubeta de cuarzo debido a su transparencia en el rango UVA, permitiendo así obtener mediciones precisas en todo espectro seleccionado. Además, se realizaron diluciones adecuadas de las muestras siguiendo la ley de Lambert-Beer, la cual establece que la absorbancia es proporcional a la concentración de la muestra y a la longitud del camino óptico. Estas diluciones fueron necesarias para mantener las lecturas dentro del rango lineal de la ley de Lambert-Beer, evitando así la saturación de la señal y asegurando una relación lineal entre la absorbancia y la concentración. De esta manera, se obtuvieron resultados cuantitativos precisos y confiables, con el fin de generar un conjunto de datos robusto que abarcara diferentes estados de crecimiento en el cultivo.

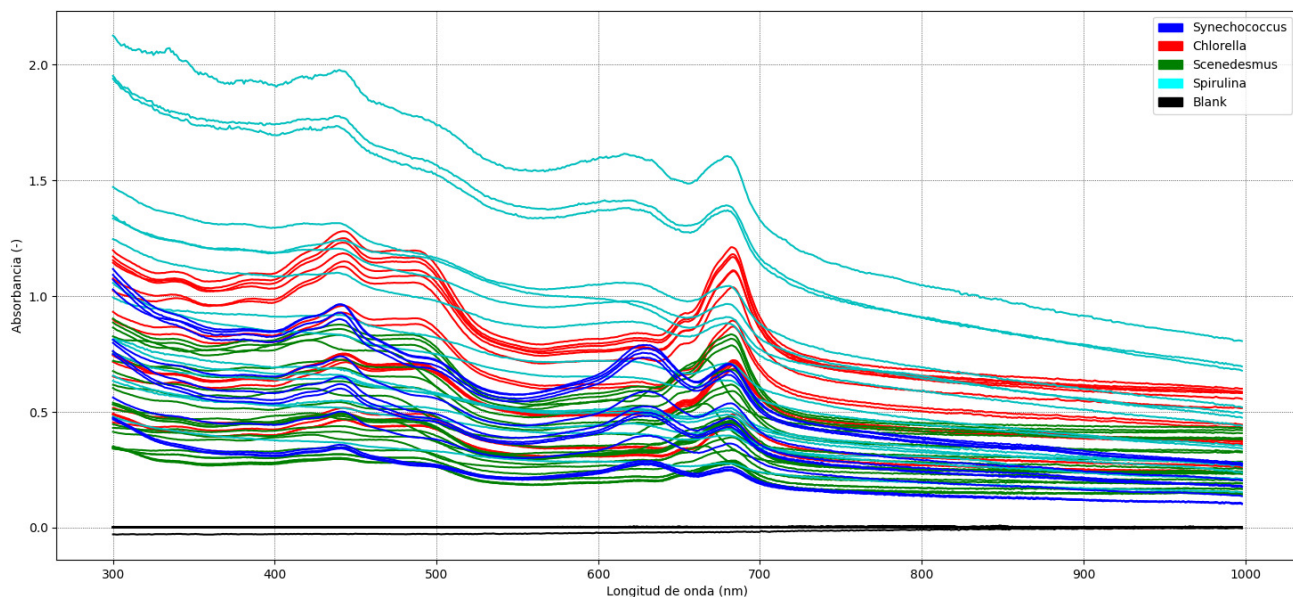


Figura 3: Barridos espectrales de los distintos géneros

Con el fin de poder crear un conjunto de datos homogéneo para el entrenamiento y validación de la red neuronal, se generó un *DataFrame*, haciendo uso de la librería *Pandas* de *Python*, en el cual se introdujo el género de la especie analizada y la absorbancia en las distintas longitudes de ondas en forma de fila. Una vez obtenido el conjunto de datos, se realizó un primer análisis visual de los barridos, con el fin de observar si los géneros presentaban patrones característicos a simple vista. En la Figura 3 se puede observar una gráfica en la cual se han muestran los 76 barridos, entre los que se incluye los barridos en blanco utilizados para la calibración del espectrofotómetro, que conforman el conjunto de datos. En ella, se observa que cada especie presenta un patrón de absorbancia diferente y por tanto, el uso de redes convolucionales de 1 dimensión puede ser útil en la extracción de información para su clasificación.

Una vez obtenido el conjunto de datos, se utilizó la biblioteca *Keras* para crear la red neuronal, la cual permitió personalizar la red definiendo las capas y funciones de activación. Se realizó una red secuencial (Figura 4) con diversas capas. Las dos primeras son capas convolucionales 1D, en las cuales se aplicó *Batch Normalization* antes de ejecutar la función de activación *ReLU* (Rectified Linear Unit), con el fin de extraer rasgos característicos de los barridos espectrales. A continuación, se introdujo una capa densa con función de activación *ReLU* encargada de interpretar los rasgos obtenidos en las capas convoluciones, seguida de una capa de *Dropout* que evitar el sobreajuste de las conexiones en la red neuronal, apagando aleatoriamente algunas de ellas durante el entrenamiento. Finalmente, se incluye una capa densa con función de activación *Softmax* con el propósito de poder realizar clasificación multi-clase.

Durante todo el proceso de diseño de la red se utilizó validación cruzada para evaluar el rendimiento de la misma de manera más exhaustiva, permitiendo mejorar su desempeño y evitando el sobreajuste. La configuración del proceso de vali-

dación cruzada fue de 10 particiones y para la arquitectura final se obtuvo una precisión media de 98.75 % con una desviación estándar de 3.75 %.

Layer (type)	Output Shape	Param #
Conv1D_1 (Conv1D)	(None, 799, 32)	128
Norm_1 (BatchNormalization)	(None, 799, 32)	128
ReLU_1 (Activation)	(None, 799, 32)	0
MaxPooling_1 (MaxPooling1D)	(None, 399, 32)	0
Conv1D_2 (Conv1D)	(None, 397, 64)	6208
Norm_2 (BatchNormalization)	(None, 397, 64)	256
ReLU_2 (Activation)	(None, 397, 64)	0
MaxPooling_2 (MaxPooling1D)	(None, 198, 64)	0
Flatten (Flatten)	(None, 12672)	0
Dropout (Dropout)	(None, 12672)	0
Dense (Dense)	(None, 64)	811072
Output (Dense)	(None, 5)	325
=====		
Total params: 818,117		
Trainable params: 817,925		
Non-trainable params: 192		

Figura 4: Arquitectura de la red

Finalmente, el modelo de red neuronal fue compilado utilizando la función de pérdida *categorical_crossentropy*, que es ampliamente utilizada en problemas de clasificación con múltiples clases. El optimizador utilizado fue SGD (*Stochastic Gradient Descent*), que es un algoritmo popular en problemas de reconcomiendo de patrones, encargado de ajustar los pesos de la red neuronal durante el entrenamiento (ichi Amari, 1993).

La métrica utilizada para evaluar el rendimiento del modelo fue la precisión (*accuracy*), que mide la proporción de predicciones correctas en relación con el total de muestras.

El conjunto de datos fue dividido en entrenamiento y validación, 70 % y 30 % respectivamente. La columna que indicaba el género fue codificada mediante la técnica de *One-hot*, esta técnica convierte cada categoría en una nueva columna binaria, asignando el valor 1 a la columna correspondiente si la muestra pertenece a esa categoría y 0 en caso contrario (Potdar et al., 2017).

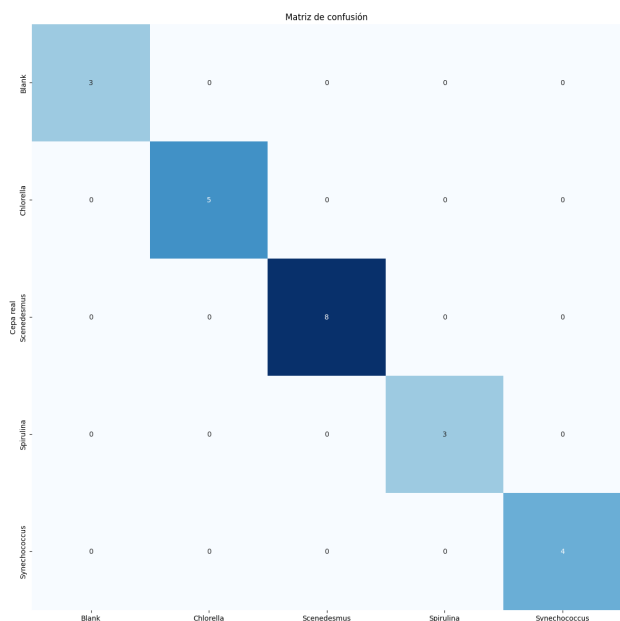


Figura 5: Matriz de confusión de validación

REAL	PREDICHO	OK?	SOFTMAX OUTPUT
Chorella	Chorella	True	[0. 100. 0. 0. 0.]
Scenedesmus	Scenedesmus	True	[0. 0. 100. 0. 0.]
Blank	Blank	True	[100. 0. 0. 0. 0.]
Blank	Blank	True	[100. 0. 0. 0. 0.]
Scenedesmus	Scenedesmus	True	[0. 0.02 99.98 0. 0.]
Chorella	Chorella	True	[0. 100. 0. 0. 0.]
Synechococcus	Synechococcus	True	[0. 0. 0. 0.02 99.98]
Blank	Blank	True	[100. 0. 0. 0. 0.]
Scenedesmus	Scenedesmus	True	[0. 2.96 97. 0.02 0.02]
Blank	Blank	True	[100. 0. 0. 0. 0.]
Synechococcus	Synechococcus	True	[0. 0. 0. 0.1 99.9]
Chorella	Chorella	True	[0. 99.93 0.07 0. 0.]
Scenedesmus	Scenedesmus	True	[0. 0.02 99.98 0. 0.]
Synechococcus	Synechococcus	True	[0. 0. 0. 0. 100.]
Scenedesmus	Scenedesmus	True	[0. 0.02 99.97 0. 0.01]
Synechococcus	Synechococcus	True	[0. 0. 0. 0.22 99.78]
Scenedesmus	Scenedesmus	True	[0. 0.03 99.97 0. 0.]
Chorella	Chorella	True	[0. 100. 0. 0. 0.]
Scenedesmus	Scenedesmus	True	[0. 0. 100. 0. 0.]
Scenedesmus	Scenedesmus	True	[0. 0.04 99.96 0. 0.]
Spirulina	Spirulina	True	[0. 0. 0.21 99.75 0.04]
Chorella	Chorella	True	[0. 99.98 0.02 0. 0.]
Synechococcus	Synechococcus	True	[0. 0. 0. 3.59 96.4]

Resultados: 23/23

Figura 6: Resultados de validación

Por último, la red neuronal fue entrenada y validada utilizando el conjunto de datos. Durante la validación, se observó que la red logró clasificar correctamente todas las muestras del conjunto de datos de validación como se muestra en la matriz de confusión (Figura 5). Además, la capa *softmax* mostró una

identificación clara de los géneros, con porcentajes de probabilidad superiores al 97 % tal y como observa en la figura 6. Estos resultados demuestran que la red neuronal aprendió efectivamente a reconocer y distinguir las características relevantes en los barridos espectrales, y fue capaz de realizar una clasificación precisa de los géneros invirtiendo un tiempo medio de 15 milisegundo en cada una de ellas.

4. Conclusiones

En conclusión, los resultados obtenidos en la clasificación de géneros utilizando espectrometrías y redes convolucionales de 1D han demostrado ser exitosos, como se ha evidenciado anteriormente. El enfoque ha permitido una clasificación precisa y eficiente de las muestras puras. Sin embargo, para futuros trabajos, se identifica la necesidad de mejorar la capa densa y la salida de la función *softmax*, con el objetivo de abordar el desafío de identificar muestras con mezclas de distintos géneros.

La identificación de muestras con mezclas es un desafío importante en el campo, ya que las muestras pueden contener diferentes géneros en proporciones variables. Mediante mejoras en la capa densa y la salida de la función *softmax*, se espera lograr una mayor capacidad de discernimiento y precisión en la identificación de estas mezclas. Esto contribuirá significativamente a la comprensión y la respuesta efectiva en situaciones donde la presencia de diferentes géneros es relevante, como la contaminación de reactores en producción a gran escala.

Agradecimientos

Este trabajo ha sido parcialmente financiado por los siguientes proyectos [números de subvención PID2020-112709RB-C21] financiados por el Ministerio de Ciencia de España y Horizonte Europa - Programa Marco de Investigación e Innovación (2021-2027) [números de subvención 101060991 REALM].

Referencias

Abdoli, S., Cardinal, P., Lameiras Koerich, A., 2019. End-to-end environmental sound classification using a 1d convolutional neural network. *Expert Systems with Applications* 136, 252–263.
 URL: <https://www.sciencedirect.com/science/article/pii/S0957417419304403>
 DOI: <https://doi.org/10.1016/j.eswa.2019.06.040>

Ación, F., Fernández, J., Magán, J., Molina, E., 2012. Production cost of a real microalgae production plant and strategies to reduce it. *Biotechnology Advances* 30 (6), 1344–1353, special issue on ACB 2011.
 URL: <https://www.sciencedirect.com/science/article/pii/S0734975012000420>
 DOI: <https://doi.org/10.1016/j.biotechadv.2012.02.005>

Anguita, D., Ghelardoni, L., Ghio, A., Oneto, L., Ridella, S., 2012. The 'k' in k-fold cross validation. In: *ESANN*, pp. 441–446.

Banerjee, S., Ramaswamy, S., 09 2017. Dynamic process model and economic analysis of microalgae cultivation in open raceway ponds. *Algal Research* 26, 330–340.
 DOI: [10.1016/j.algal.2017.08.011](https://doi.org/10.1016/j.algal.2017.08.011)

Bargchi, H., Dehnavi, Z., Nattagh-Eshtivani, E., Alwaily, E. R., Almulla, A. F., Kareem, A. K., Barati, M., Ranjbar, G., Mohammadzadeh, A., Rahimi, P., Pahlavani, N., 2023. The effects of chlorella vulgaris on cardiovascular risk factors: A comprehensive review on putative molecular mechanisms. *Biomedicine Pharmacotherapy* 162, 114624.
 URL: <https://www.sciencedirect.com/science/article/pii/S0753332223004122>
 DOI: <https://doi.org/10.1016/j.biopha.2023.114624>

- Guzmán, J. L., Acién, F. G., Berenguel, M., 12 2021. Modelling and control of microalgae production in industrial photobioreactors. *Revista Iberoamericana de Automática e Informática Industrial* 18 (1), 1–18.
DOI: 10.4995/riai.2020.13604
- ichi Amari, S., 1993. Backpropagation and stochastic gradient descent method. *Neurocomputing* 5 (4), 185–196.
URL: <https://www.sciencedirect.com/science/article/pii/0925231293900060>
DOI: [https://doi.org/10.1016/0925-2312\(93\)90006-0](https://doi.org/10.1016/0925-2312(93)90006-0)
- Kim, Y., Oct. 2014. Convolutional neural networks for sentence classification. In: *Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP)*. Association for Computational Linguistics, Doha, Qatar, pp. 1746–1751.
URL: <https://aclanthology.org/D14-1181>
DOI: 10.3115/v1/D14-1181
- Kim, Y., Jeon, J., Kwak, M., Kim, G. H., Koh, I., Rho, M., 01 2018. Photosynthetic functions of *Synechococcus* in the ocean microbiomes of diverse salinity and seasons. *PLOS ONE* 13, e0190266.
DOI: 10.1371/journal.pone.0190266
- Ko, C. W., Huh, J., Park, J.-W., 2022. Deep learning program to predict protein functions based on sequence information. *MethodsX* 9, 101622.
URL: <https://www.sciencedirect.com/science/article/pii/S2215016122000073>
DOI: <https://doi.org/10.1016/j.mex.2022.101622>
- Kozlova, T., Hardy, B., Krishna, P., Levin, D., 10 2017. Effect of phytohormones on growth and accumulation of pigments and fatty acids in the microalgae *scenedesmus quadricauda*. *Algal Research* 27, 325–334.
DOI: 10.1016/j.algal.2017.09.020
- Kumar, A., Ramamoorthy, D., Verma, D. K., Kumar, A., Kumar, N., Kanak, K. R., Marwein, B. M., Mohan, K., 2022. Antioxidant and phytonutrient activities of *spirulina platensis*. *Energy Nexus* 6, 100070.
URL: <https://www.sciencedirect.com/science/article/pii/S277242712200033X>
DOI: <https://doi.org/10.1016/j.nexus.2022.100070>
- Kumaran, M., Palanisamy, K. M., Bhuyar, P., Maniam, G. P., Rahim, M. H. A., Govindan, N., 2023. Agriculture of microalgae *chlorella vulgaris* for polyunsaturated fatty acids (pufas) production employing palm oil mill effluents (pome) for future food, wastewater, and energy nexus. *Energy Nexus* 9, 100169.
URL: <https://www.sciencedirect.com/science/article/pii/S2772427122001243>
DOI: <https://doi.org/10.1016/j.nexus.2022.100169>
- Nordio, R., Delgado, F., Sánchez-Zurano, A., Gonzalez Hernandez, J., Rodriguez Miranda, E., Guzmán, J. L., Lafarga, T., Acién, G., 11 2022. Long-term assessment of the nutrient recovery capacity and biomass productivity of *scenedesmus almeriensis* in raceway reactors using unprocessed urban wastewater. *Bioresource Technology* 369, 128374.
DOI: 10.1016/j.biortech.2022.128374
- Otálora Berenguel, P., Guzmán, J., Acién, G., Berenguel, M., Reul, A., 05 2021. Microalgae classification based on machine learning techniques. *Algal Research* 55, 102256.
DOI: 10.1016/j.algal.2021.102256
- Park, J., Craggs, R., Shilton, A., 2011. Wastewater treatment high rate algal ponds for biofuel production. *Bioresource Technology* 102 (1), 35–42, special Issue: Biofuels - II: Algal Biofuels and Microbial Fuel Cells.
URL: <https://www.sciencedirect.com/science/article/pii/S0960852410011636>
DOI: <https://doi.org/10.1016/j.biortech.2010.06.158>
- Potdar, K., Pardawala, T., Pai, C., 10 2017. A comparative study of categorical variable encoding techniques for neural network classifiers. *International Journal of Computer Applications* 175, 7–9.
DOI: 10.5120/ijca2017915495
- Roy, C. J., Khoo, K. S., Kit Wayne, C., Vo, D.-V., Deepanraj, B., Banat, F., Munawaroh, H., Iwamoto, K., Show, P.-L., 12 2022. Use of image processing and digital algorithm for microalgae identification. *Bioresource Technology* 369, 128418.
DOI: 10.1016/j.biortech.2022.128418
- Spolaore, P., Joannis-Cassan, C., Duran, E., Isambert, A., 2006. Commercial applications of microalgae. *Journal of Bioscience and Bioengineering* 101 (2), 87–96.
URL: <https://www.sciencedirect.com/science/article/pii/S1389172306705497>
DOI: <https://doi.org/10.1263/jbb.101.87>
- Tapie, P., Bernard, A., 1988. Microalgae production: Technical and economic evaluations. *Biotechnology and Bioengineering* 32 (7), 873–885.
URL: <https://onlinelibrary.wiley.com/doi/abs/10.1002/bit.260320705>
DOI: <https://doi.org/10.1002/bit.260320705>
- Villalobos, M. B., Fernández, F. G. A., Guzmán, J. L., Sevilla, J. M. F., Berenguel, M., 2019. New strategies for the design and control of raceway reactors to optimize microalgae production. *Handbook of Algal Technologies and Phytochemicals*.
- Zhao, B., Su, Y., 2014. Process effect of microalgal-carbon dioxide fixation and biomass production: A review. *Renewable and Sustainable Energy Reviews* 31, 121–132.
URL: <https://www.sciencedirect.com/science/article/pii/S1364032113008034>
DOI: <https://doi.org/10.1016/j.rser.2013.11.054>