

# Enfoque Híbrido para la Clasificación de Muestras Histológicas para la Detección de Tumores Cerebrales: Descomposición lineal y Redes Neuronales

I. A. Cruz-Guerrero<sup>1\*</sup>, D. U. Campos-Delgado<sup>1,2</sup>, A. R. Mejía-Rodríguez<sup>1</sup>, S. Ortega<sup>3</sup>, H. Fabelo<sup>3</sup>, G. M. Callico<sup>3</sup>

<sup>1</sup>Facultad de Ciencias, Universidad Autónoma de San Luis Potosí, San Luis Potosí, México.

<sup>2</sup>Instituto de Investigación en Comunicación Óptica, Universidad Autónoma de San Luis Potosí, San Luis Potosí, México.

<sup>3</sup>Instituto Universitario de Microelectrónica Aplicada, Universidad de las Palmas de Gran Canaria, Las Palmas de Gran Canaria, España.

\*a211439@alumnos.uaslp.mx

**Abstract**— Hyperspectral imaging has demonstrated its potential to provide information of the chemical composition of tissue and also of its morphological characteristics. However, discerning the presence of a pathology through this information is not a simple task. Because of this, a hybrid methodology is proposed in this work, which combines the identification of characteristic components present in a hyperspectral image from linear unmixing methods, and the ability to distinguish patterns from a neural network. The results of this research show that the proposed method can distinguish a tumor condition from histological brain samples with an average accuracy of 87%. The study demonstrates the potential of hybrid classification methodologies in the analysis of spectral information for the identification of histological samples affected by tumor tissue.

**Palabras clave**— descomposición lineal, imágenes hiperespectrales, muestras histológicas, redes neuronales, tumor cerebral.

## I. INTRODUCCIÓN

El diagnóstico tradicional en histopatología se basa en realizar un análisis manual de las características morfológicas presentes en las muestras. No obstante, en los últimos años el uso y aplicación de nuevas tecnologías en estudios histológicos ha impulsado esta tarea, mejorando el diagnóstico de muestras patológicas para aumentar la objetividad, disminuir el tiempo del examen y reducir la variabilidad entre observadores [1-2]. Una de las principales herramientas que han surgido para mejorar los estudios patológicos es el análisis de imágenes RGB (rojo, verde y azul) de muestras histológicas [1]. No obstante, en los últimos años han surgido nuevas herramientas que permiten adquirir una mayor cantidad de información, tales como las imágenes hiperespectrales (HSI, por sus siglas en inglés), las cuales aportan información tanto espacial como espectral de las muestras a analizar [3-4]. Recientemente, esta tecnología ha demostrado ser de utilidad en múltiples campos de aplicación, como la visualización de marcadores biológicos dentro de una muestra de tejido con inmunohistoquímica, la tinción digital de muestras o el diagnóstico de diversos tipos de patologías [5].

A pesar de las ventajas que aporta el considerar la información hiperespectral (HS, por sus siglas en inglés), el

distinguir y clasificar este tipo de información no es una tarea sencilla. Debido a esto, en la literatura se han propuesto diversas metodologías que permiten identificar y clasificar los componentes presentes en una HSI, siendo dos de las más utilizadas las técnicas de descomposición lineal [6] y estrategias de aprendizaje de máquina [7]. Los métodos de descomposición lineal permiten identificar los componentes característicos y la proporción de estos en una HSI, pero son deficientes ante la presencia de difracciones e interacciones no lineales [6]. Por su parte, las técnicas de aprendizaje de máquina permiten realizar la identificación y clasificación de información mediante el reconocimiento de rasgos o patrones principalmente de manera supervisada, no obstante, es necesario una gran cantidad de datos de entrenamiento para producir resultados aceptables [7].

Debido a esto, en este trabajo se propone una metodología híbrida que combina la identificación de componentes característicos presentes en una HSI de los métodos de descomposición lineal, y la habilidad de distinguir patrones del aprendizaje de máquina. En específico, se emplea una red neuronal multicapa para generar un algoritmo que permite diferenciar la presencia de tejido tumoral en muestras histológicas.

## II. METODOLOGÍA

### A. Descripción de base de datos

La base de datos para este trabajo consistió en un conjunto de HSI adquiridas a partir de muestras histológicas de cerebro humano, tal y como se describe en [8-9]. Estas muestras fueron procesadas y analizadas por el Departamento de Anatomía Patológica del Hospital Universitario Doctor Negrín de Las Palmas de Gran Canaria (España), cuyo Comité de Investigación Clínica-Comité de Ética en la Investigación (CEIC/CEI) proveyó su consentimiento y aprobó el protocolo de estudio. El diagnóstico definitivo de las muestras fue proporcionado por parte de patólogos de la institución antes mencionada, determinando el diagnóstico del tejido según la clasificación de tumores del sistema nervioso de la Organización Mundial de la Salud (OMS) [10]. Las HSIs se capturaron utilizando un microscopio equipado con una cámara HS de barrido, adquiridas con una magnificación de 20×, produciendo un tamaño de imagen HS

de  $375 \times 299 \mu\text{m}$ . En la FIG. 1 se observa un ejemplo de las imágenes RGB sintéticas de microscopía (generadas a partir de la HSI) para las dos clases de tejido estudiadas. El rango espectral de las imágenes va de 400 a 1000 nm con una resolución espectral de 2.8 nm, generando HSIs de  $1004 \times 800$  píxeles con 826 canales espectrales. Después de una etapa de preprocesamiento descrita en [8], los cubos HS resultantes tienen una resolución de 275 canales espectrales.

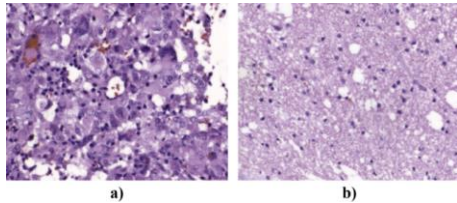


Fig. 1. Ejemplo de imágenes RGB sintéticas de microscopía capturadas con una magnificación de 20x: a) muestra tumoral, b) muestra no tumoral.

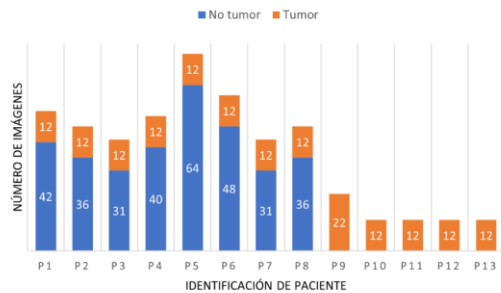


Fig. 2. Número de HSI de tumor y no tumor por paciente.

La base de datos total cuenta con 494 HSIs adquiridas de 13 muestras histológicas de 13 pacientes diferentes con tumor tipo Glioblastoma (grado IV), cuya distribución de imágenes por paciente se muestra en la FIG. 2. Como se puede apreciar, los pacientes poseen diferente número de imágenes etiquetadas como no tumoral. Esta diferenciación en el conjunto de datos plantea tres problemas: el número limitado de pacientes, las muestras que contienen ambas clases (no tumorales y tumorales) sólo están disponibles para ocho pacientes, y finalmente el conjunto de datos no está equilibrado, ya que tiene más imágenes etiquetadas como no tumorales. Debido a esto, se optó por realizar una serie de particiones para realizar la evaluación del modelo híbrido, esto con el fin de evitar problemas de sesgos y sobreajuste al momento de realizar la etapa de clasificación. Enseguida, se dividieron los datos en cuatro conjuntos diferentes, cada uno de ellos construido con un subconjunto de entrenamiento, validación y prueba independiente de los pacientes. Tres particiones compuestas por nueve pacientes de entrenamiento, uno de validación y tres de prueba, mientras que la partición restante consta de ocho pacientes de entrenamiento, uno de validación y cuatro de prueba. La asignación de los pacientes en cada partición se realizó de manera aleatoria procurando que cada uno se presente una vez en el subconjunto de prueba y seleccionando en el

subconjunto de validación a pacientes que posean ambos tipos de datos etiquetados (no tumorales y tumorales), destinando dos pacientes que solo poseen información tumoral al subconjunto de entrenamiento [8-9].

### B. Metodología Híbrida de Clasificación

La metodología de clasificación híbrida utilizada en este trabajo consta de dos partes principales. La primera etapa se centra en un esquema de descomposición lineal que reduce la dimensionalidad de los datos a clasificar capturando las principales características de la HSI [6, 11]. La segunda etapa consiste en un clasificador basado en redes neuronales [12].

La etapa de descomposición lineal permite obtener mapas de abundancia y miembros finales característicos (también llamadas firmas espectrales) a partir de la HSI, siendo estos últimos independientes para cada compuesto presente en el cubo HS. Por su parte, los mapas de abundancia indican la proporción de dichos componentes. La descomposición de cada una de las 494 imágenes se efectuó mediante el algoritmo de extracción ciega ampliada de miembros finales y abundancias (EBEAE por sus siglas en inglés) descrito en [11], utilizando como método de inicialización de miembros finales el algoritmo de N-FINDR [13]. Por otro lado, los parámetros de control de EBEAE se establecieron en  $\rho = 1, \lambda = 0, \epsilon = 1 \times 10^{-6}$  y un número máximo de 10 iteraciones, con un orden de búsqueda de cuatro componentes (i.e.  $N=4$ ). Una vez estimados los mapas de abundancia y miembros finales de todas las imágenes, se guardaron los datos producidos de cada imagen por paciente, para posteriormente establecer las particiones comentadas en la sección anterior. No obstante, existe un problema intrínseco de las muestras histológicas, esto es la interferencia causada por las regiones en las muestras que permiten el paso directo de la luz del microscopio, la cual no aporta información sobre los componentes presentes en el tejido. Para evitar la interferencia que podría causar la información de dichas regiones, se optó por eliminar tanto el miembro final como el mapa de abundancia correspondiente a las zonas que no atraviesan muestras de tejido, es decir el asociado a la luz blanca.

La segunda etapa consiste en un ensamble de redes neuronales, esto con el fin de disminuir la variabilidad de los resultados y mejorar la exactitud general de la metodología propuesta [12]. El ensamble se realizó mediante la unión de dos redes neuronales de nivel 0, las cuales son idénticas y constan de cinco capas internas con función de activación *LeakyReLU* con parámetro  $\alpha = 0.1$  y una etapa de *dropout* del 50%, con función de activación *softmax* para el caso de la capa de salida. Para la etapa de entrenamiento, se empleó la función de costo de entropía binaria cruzada [14]. Los datos de entrada de la red consisten en los miembros finales concatenados de acuerdo con la similitud entre ellos, es decir, en la primera posición se colocó la firma espectral que posee mayor similitud con respecto a los miembros finales del resto de HSI, mientras que al final las que poseen mayor diferencia. La similitud se evaluó con respecto del error cuadrático

medio entre los miembros finales. Al terminar este proceso se obtiene la red entrenada, la cual es guardada para su utilización en la red ensamblada o red de nivel 1.

La red de nivel 1 utiliza como base las dos redes de nivel 0 preentrenadas, sin considerar la capa de salida de estas últimas, al concatenar las salidas de ambas redes y formando un vector de 40 elementos. Este vector entra a una red formada por dos capas internas de 100 y 32 neuronas con función de activación *LeakyRelu* con  $(\alpha = 0.1)$  en ambas capas y una etapa de *dropout* del 50% en la primera. De igual manera que en el caso de las redes de nivel 0, la función de costo durante el entrenamiento fue entropía binaria cruzada y con *softmax* como función de activación de la capa de salida. Finalmente, la salida de la red devuelve un valor binario, cero para la etiqueta de No Tumor y uno para Tumor, lo cual se verifica con el etiquetado realizado por el experto clínico, presentando la arquitectura general de la red en la FIG. 3.

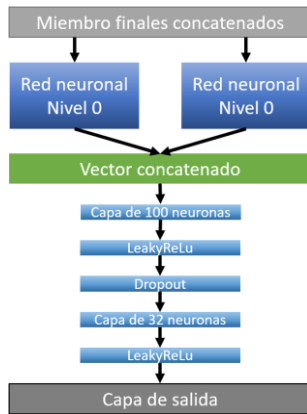


Fig. 3. Arquitectura de la red neuronal multicapa de nivel 1.

### C. Métricas de Clasificación

La evaluación de la metodología propuesta se realizó mediante el análisis de diversas métricas de clasificación ampliamente utilizadas en el estado del arte, tal como la exactitud, la sensibilidad, la precisión y F1-score [15].

## III. RESULTADOS

En esta sección se presentan los resultados de la clasificación generada por el método híbrido. Para validar los resultados producidos se utilizó una metodología de superpíxeles en conjunto con una máquina de soporte vectorial (SVM por sus siglas en inglés) cuyos parámetros fueron optimizados mediante balanceo de exactitud, esta metodología es descrita en [9]. El entrenamiento de la red ensamblada se efectuó con el conjunto de datos asignado de cada partición. Estos datos también se usaron en el entrenamiento de las redes de nivel 0. En la etapa de entrenamiento de la red de nivel 1 la exactitud promedio de las diversas particiones fue de un 95%, no obstante, esta forma de entrenamiento es susceptible a sobreajuste [12]. Por lo que la comprobación y clasificación se realizó mediante los datos de validación y de entrenamiento de cada partición,

evaluando el nivel de sobreajuste del clasificador. Debido a esto, en este trabajo nos centramos en los datos de validación y prueba de las diferentes particiones. La implementación del método propuesto se realizó mediante dos plataformas diferentes. La etapa de preprocesamiento y descomposición se implementó en el software Matlab. Mientras que la etapa de clasificación y evaluación de las predicciones se llevó a cabo en el lenguaje de programación Python, esto a causa de las librerías y herramientas que ofrece cada entorno.

### A. Resultados del Conjunto de Validación

La FIG. 4 presenta los resultados producidos por el método híbrido y la técnica de superpíxeles con SVM, en el conjunto de datos de validación. Observando diferencias significativas en la partición 1, donde las métricas se distancian en más de un 7% entre ambas metodologías. Por otro lado, el resto de las particiones siguen una tendencia similar en ambos métodos. Además, los valores promedio de cada métrica resulta en exactitud  $0.86 \pm 0.13$  (valor medio  $\pm$  desviación estándar), sensibilidad  $0.75 \pm 0.40$ , precisión  $0.70 \pm 0.30$  y F1-score  $0.69 \pm 0.33$  para el caso del método híbrido. Mientras que en el caso de superpíxeles con SVM  $0.88 \pm 0.16$ ,  $0.87 \pm 0.20$ ,  $0.86 \pm 0.28$  y  $0.83 \pm 0.18$  para las mismas métricas respectivamente.

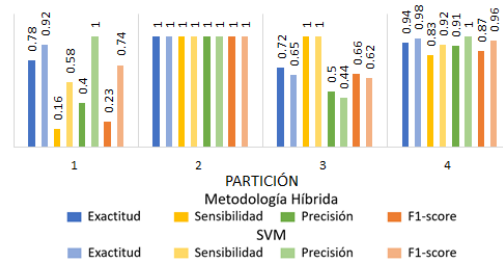


Fig. 4. Resultados de clasificación en el conjunto de datos de validación para cada partición propuesta, en comparación con los resultados de superpíxeles con SVM [9].

### B. Resultados del Conjunto de Prueba

Los resultados de la evaluación en los datos de prueba se presentan en la FIG. 5, mostrando en primera instancia un aumento en el desempeño de todas las métricas, al comparar con los resultados de validación (ver FIG. 4). Apreciando un menor desempeño en la partición 2, mientras que en el caso de las particiones 1, 3 y 4 se produjeron mejores resultados con respecto al caso de validación. Los promedios producidos

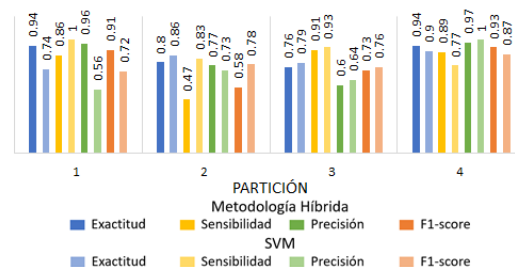


Fig. 5. Resultados de clasificación en el conjunto de datos de prueba para cada partición propuesta, en comparación con los resultados de superpíxeles con SVM [9].

son de  $0.87 \pm 0.09$ ,  $0.78 \pm 0.21$ ,  $0.83 \pm 0.18$  y  $0.79 \pm 0.16$  para las métricas de exactitud, sensibilidad, precisión y F1-score en el caso del método híbrido; mientras que, en el caso de superpíxeles con SVM se obtuvo  $0.82 \pm 0.07$ ,  $0.88 \pm 0.10$ ,  $0.73 \pm 0.19$  y  $0.78 \pm 0.06$ , respectivamente. Estos resultados nos permiten comprobar que el entrenamiento presenta un bajo o nulo nivel de sobreajuste en el entrenamiento.

#### IV. DISCUSIÓN

Los resultados presentados en la FIG. 4 y FIG. 5 nos permiten observar el comportamiento de la metodología propuesta en distintos conjuntos de datos. A pesar del buen desempeño presentado en los datos de validación, no es posible asegurar que no se presenta un sobreajuste en el entrenamiento solo con los datos de un paciente. Sin embargo, al analizar a la par los resultados del conjunto de datos de prueba se aprecia un rendimiento incluso superior en las distintas particiones propuestas. Esto implica que a pesar de entrenar las redes de nivel 0 y nivel 1 con los mismos datos de entrenamiento, la arquitectura propuesta limita el impacto de sobreajuste en el entrenamiento. Los resultados producidos son comparables a lo publicado en el estado del arte [9], donde se produce una exactitud promedio de 88.7% y 82.6% para los conjuntos de validación y prueba, mientras que en el caso de la propuesta se genera una exactitud promedio de 86.2% y 86.5% para los mismos datos, demostrando estabilidad en las estimaciones para los diversos conjuntos, además de presentar un aumento de 4% en la exactitud de los datos de prueba. Además de esto, el tiempo de entrenamiento de la propuesta es de aproximadamente 5 minutos, en comparación con la SVM que demoró alrededor de 4 horas en entrenarse con los mismos datos. Aún considerando la etapa de preprocesamiento y descomposición de todas las HSIs la propuesta tarda menos de aproximadamente 40 minutos en realizar todo el proceso, lo que representa una reducción del 83% del tiempo con un ligero aumento en las métricas de evaluación.

Pese a los buenos resultados producidos, una de las principales limitaciones de este estudio es el reducido número de pacientes y el desequilibrio del conjunto de datos (ver FIG. 2). Por lo tanto, es necesario contar con más pacientes para obtener conclusiones más sólidas sobre la capacidad de las metodologías híbridas para la detección automática de glioblastoma en cortes histológicos.

#### V. CONCLUSIONES

En este estudio, demostramos el potencial de las metodologías híbridas de clasificación HS. Nuestra propuesta conjuga algoritmos de descomposición lineal de HSI con métodos de clasificación como redes neuronales, en el análisis de información espectral para la identificación de muestras histológicas afectadas por tejido tumoral. Los resultados comprueban la eficacia del método propuesto, para la clasificación de muestras histológicas de tejido, produciendo un nivel de exactitud remarcable en las imágenes

analizadas. Además de esto, la metodología propuesta disminuye de manera significativa el costo computacional, lo que se traduce en un menor tiempo de entrenamiento sin comprometer el rendimiento de la clasificación.

#### RECONOCIMIENTO

Este trabajo fue apoyado por el CONACYT a través de un proyecto de Ciencia Básica (Ref. #254637). El autor I. A. Cruz-Guerrero agradece al CONACYT por su apoyo a través de una beca de doctorado (Ref. #865747).

#### DECLARACIONES ÉTICAS

El protocolo del estudio y los consentimientos fueron aprobados por el Comité Ético de Investigación Clínica-Comité de Ética en la Investigación (CEIC/CEI) del Hospital Universitario Doctor Negrín: Identificación Hiperespectral de tumores Cerebrales (ITHaCA), Código: 2019-001-1.

#### REFERENCIAS

- [1] S. L. Van Es, "Digital pathology: semper ad meliora," *Pathology*, vol. 51, no. 1, pp. 1–10, 2019.
- [2] T. J. Flotte and D. A. Bell, "Anatomical pathology is at a crossroads," *Pathology*, vol. 50, no. 4, pp. 373–374, 2018.
- [3] T. H. Johansen *et al.*, "Recent advances in hyperspectral imaging for melanoma detection," *Wiley Interdiscip. Rev. Comput. Stat.*, vol. 12, no. 1, p. e1465, 2020.
- [4] M. Halicek *et al.* "In-vivo and ex-vivo tissue analysis through hyperspectral imaging techniques: revealing the invisible features of cancer," *Cancers (Basel)*, vol. 11, no. 6, p. 756, 2019.
- [5] T. H. Johansen *et al.*, "Recent advances in hyperspectral imaging for melanoma detection," *Wiley Interdiscip. Rev. Comput. Stat.*, vol. 12, no. 1, p. e1465, 2020.
- [6] J. M. P. Nascimento and J. M. Bioucas-Dias, "Nonlinear mixture model for hyperspectral unmixing," in *Image and Signal Processing for Remote Sensing XV*, 2009, vol. 7477, no. September 2009, p. 74770I, doi: 10.1117/12.830492.
- [7] U. B. Gwali, S. T. Monteiro, and E. Saber, "Machine learning based hyperspectral image analysis: a survey," *arXiv Prepr. arXiv1802.08701*, 2018.
- [8] Ortega, Samuel, *et al.* "Hyperspectral imaging for the detection of glioblastoma tumor cells in H&E slides using convolutional neural networks". *Sensors*, 2020, vol. 20, no. 7, p. 1911.
- [9] Ortega, Samuel, *et al.* "Hyperspectral superpixel-wise glioblastoma tumor detection in histological samples". *Applied Sciences*, 2020, vol. 10, no. 13, p. 4448.
- [10] Louis, David N., *et al.* "The 2016 World Health Organization classification of tumors of the central nervous system: a summary. *Acta neuropathologica*", 2016, vol. 131, no. 6, p. 803-820.
- [11] D. U. Campos-Delgado *et al.*, "Extended Blind End-Member and Abundance Extraction for Biomedical Imaging Applications," *IEEE Access*, vol. 7, pp. 178539–178552, 2019, doi: 10.1109/ACCESS.2019.2958985.
- [12] J. Friedman, T. Hastie, and R. Tibshirani, "The elements of statistical learning", 2nd ed, New York, USA, Springer series in statistics, 2001, ch 11 and 17.
- [13] M. E. Winter, "N-FINDR: An algorithm for fast autonomous spectral end-member determination in hyperspectral data," in *Imaging Spectrometry V*, 1999, vol. 3753, pp. 266–275.
- [14] Y. Ho and S. Wookey, "The real-world-weight cross-entropy loss function: Modeling the costs of mislabeling," *IEEE Access*, vol. 8, pp. 4806–4813, 2019.
- [15] Hossin, Mohammad; Sulaiman, Md Nasir. "A review on evaluation metrics for data classification evaluations". *International journal of data mining & knowledge management process*, 2015, vol. 5, no. 2, p. 1.