

CLASIFICACIÓN DE TUMORES CEREBRALES EN IMÁGENES DE RESONANCIA MAGNÉTICA MEDIANTE PARTICIONAMIENTO DE REGIONES Y BÚSQUEDA HEURÍSTICA

*MAGNETIC RESONANCE IMAGE BASED BRAIN TUMOR
CLASSIFICATION THROUGH REGION PARTITIONING AND
HEURISTIC SEARCH*

Néstor Uriel Hernández Cortez

Universidad Tecnológica de la Mixteca, México
hecn980226@gs.utm.mx

Luis Enrique Reyes Martínez

Universidad Tecnológica de la Mixteca, México
rem1981220@gs.utm.mx

Jazmin Torres Bautista

Universidad Tecnológica de la Mixteca, México
tobj971105@gs.utm.mx

Joseph Velazquez Morales

Universidad Tecnológica de la Mixteca, México
venj980903@gs.utm.mx

Francisco Isaac Reyes Sánchez

Universidad Tecnológica de la Mixteca, México
resf911010@gs.utm.mx

Raúl Cruz Barbosa

Universidad Tecnológica de la Mixteca, México
rcruz@gs.utm.mx

Recepción: 3/noviembre/2021

Aceptación: 7/marzo/2022

Resumen

En este trabajo se presenta un clasificador tradicional de tres tipos de tumores cerebrales (meningioma, glioma y pituitario) basado en imágenes de resonancia magnética. Para conseguirlo se extrajeron características de intensidad, textura y forma a la región tumoral de interés (ROI), así como a la ROI aumentada y particionada en anillos; posteriormente los vectores de características fueron

enviados a selectores de características como el índice discriminante de Fisher, métodos de empaquetamiento y un algoritmo genético para comparar el rendimiento con distintos clasificadores tradicionales. Los resultados obtenidos muestran que el uso de un algoritmo genético para la selección de características y búsqueda de parámetros brinda un mejor rendimiento, permitiendo mejorar la clasificación reportada en la literatura utilizando el mismo conjunto de datos y clasificadores tradicionales. Es decir, se obtiene una exactitud de clasificación de 92.35% utilizando una máquina de soporte vectorial, con características de la ROI y partición en anillos.

Palabras clave: Algoritmo genético, aumento y particionamiento de región tumoral, clasificación de tumores cerebrales, imagen de resonancia magnética.

Abstract

In this work a conventional classifier of three types of brain tumors (meningioma, glioma and pituitary) based on magnetic resonance images is presented. To achieve it, features of intensity, texture and shape were extracted from the tumor region of interest (ROI), as well as from the increased ROI and partitioned into rings; They were subsequently subjected to feature selectors such as Fisher's discriminant index, wrapper methods, and a genetic algorithm to compare performance with different conventional classifiers. The obtained results show that the use of a genetic algorithm for feature selection and parameter search provides better performance, allowing to improve those reported in the literature using the same data set and conventional classifiers. A classification accuracy of 92.35% is obtained with a support vector machine, using features extracted from the ROI and the augmented and partitioned ROI.

Keywords: Brain tumor classification, genetic algorithm, magnetic resonance imaging, tumor augmented and partitioned region.

1. Introducción

Un tumor es un crecimiento anormal de tejidos en cualquier parte del cuerpo, el cual puede ser maligno o benigno. Por su parte, un tumor cerebral es más grave y

potencialmente mortal debido a que se aloja en un espacio limitado de la cavidad intracraneal, además, sus características varían dependiendo al tipo que pertenezca [Lala, 2013]. En el presente trabajo se abordan los tumores cerebrales más frecuentes, como son meningioma, glioma y tumor pituitario; el segundo es considerado maligno mientras que los dos restantes son considerados como benignos.

Existen diferentes técnicas para obtener imágenes cerebrales, las principales son la resonancia magnética, tomografía por emisión de positrones, y la tomografía computarizada. Las imágenes de resonancia magnética ponderadas en T1 se utilizan principalmente para el diagnóstico de tumores cerebrales primarios debido a que ofrece una buena caracterización de tejidos y ausencia de artefactos óseos y dentales [Angulakshmi, 2017]; [Rehman, 2020].

La clasificación de tumores cerebrales es el proceso de categorizar tumores en función de sus características. Actualmente, existen diversos estudios sobre la clasificación de tumores cerebrales, por ejemplo, en [Cheng, 2015] proponen un método que mejora la clasificación mediante el aumento de la región tumoral y su posterior partición en subregiones en forma de anillo, debido a que los tejidos que rodean al tumor también proporcionan información relevante. Por su parte en [Ismael, 2018] proponen un método basado en la extracción de características estadísticas de primer y segundo orden, características mediante la transformada discreta de Wavelet y filtros de Gabor, utilizando una red neuronal perceptrón multicapa para la clasificación. Por otro lado, en [Rehman, 2020] utilizaron tres arquitecturas de Redes Neuronales Convolucionales (RNC) como son AlexNet, GoogLeNet y VGGNet. Emplearon un aumento de datos para reducir la posibilidad de sobreajuste y concluyeron que el modelo VGGNet es el más adecuado. [Abiwinanda, 2018] propusieron cinco arquitecturas de RNC las cuáles se caracterizan por ser más simples que AlexNet, VGG16 y ResNet, encontraron que la segunda arquitectura propuesta les arrojó mejores resultados de clasificación. En [Afshar, 2018] enfocan el estudio en diseñar y evaluar la arquitectura CapsNet, la cual arroja mejores resultados de clasificación de tumores cerebrales al compararla con otras redes neuronales convolucionales.

El presente trabajo tiene como objetivo mostrar un enfoque sencillo con clasificadores tradicionales, donde la importancia radica en la selección de características relevantes y búsqueda de parámetros adecuados de dichos clasificadores a través de un algoritmo genético. Es por ello que se utilizaron características de intensidad, textura y forma, dejando a un lado características más sofisticadas. De igual manera se emplearon clasificadores tradicionales como el Bayesiano, Bayes ingenuo, k vecinos más cercanos (KNN), máquina de soporte vectorial (SVM) y bosques aleatorios (RF o random forest). Por lo mencionado anteriormente, el proceso de encontrar un subconjunto de características que proporcione una buena separabilidad de clases obtuvo un papel importante. Para ello se utilizaron tres enfoques para la selección de características, como son: selección mediante ranqueo utilizando el índice discriminante de Fisher; métodos de empaquetamiento y búsqueda heurística mediante un algoritmo genético.

2. Métodos

Previo a la metodología para clasificación, primero se obtienen las máscaras de las diferentes regiones a partir de la máscara de la región tumoral proporcionada por un experto humano e incluida en el conjunto de datos. Entonces, la metodología utilizada en este trabajo consta de las siguientes etapas: a) Extracción de características, b) Normalización, c) Selección de características, d) Clasificación. Esto es, a las imágenes originales se les extrae las características a estudiar y después se normalizan para formar los vectores que componen el conjunto original de datos. Posteriormente, se selecciona el mejor subconjunto de características que ofrezca el rendimiento de clasificación más alto. A continuación, se presentan los métodos utilizados en cada etapa de la metodología.

Extracción de características

Debido a la estructura complicada de los tejidos cerebrales, como la materia blanca, materia gris y líquido cefalorraquídeo, la extracción de características es una tarea crucial para la clasificación, ya que dichas características proporcionan una caracterización completa del tumor y se utilizan como entrada para los clasificadores

[Saman, 2019]. En el presente proyecto se extrajeron tres tipos de características básicas, como son: basadas en intensidad, textura y forma. Además, se extrajeron características adicionales obtenidas del aumento y particionamiento de la región tumoral.

Características basadas en intensidad

Las características basadas en intensidad se calculan mediante el histograma de intensidad de la imagen. La densidad de probabilidad de ocurrencia de los niveles de intensidad se obtiene dividiendo cada valor del histograma entre el número total de píxeles, como se muestra en la ecuación 1.

$$p(i) = H(i) / (n * m) \quad (1)$$

Donde H es el histograma y m, n son las dimensiones verticales y horizontales de la imagen. Se calcularon 4 características de intensidad, como son: media, varianza, asimetría y curtosis [Saman, 2019].

Características basadas en textura

Las características basadas en textura se calculan mediante la matriz de coocurrencia de nivel de gris (GLCM). Se calculó la matriz de coocurrencia simétrica en cuatro direcciones: 0° , 45° , 90° y 135° con distancia de 1. Se sumaron las cuatro matrices para formar una única matriz y posteriormente se obtuvo la densidad de probabilidad dividiendo cada valor de la matriz entre el número total de píxeles, como se muestra en la ecuación 2.

$$p(i, j) = GLCM(i, j) / (n * m) \quad (2)$$

Donde $GLCM$ es la matriz única y n, m son las dimensiones verticales y horizontales de la imagen. Se calcularon 14 características de textura, como son: autocorrelación, correlación, inercia, energía, homogeneidad, entropía, disimilaridad, forma de clúster, sombra de clúster, prominencia de clúster, momento de diferencia inversa, suma promedio, varianza de la suma, probabilidad máxima [Haralick, 1973]; [Qurat-UI-Ain, 2010]; [Zulpe, 2012]; [Saman, 2019].

Características basadas en forma

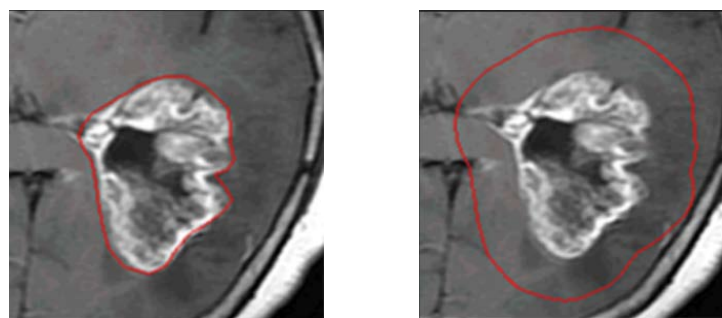
Las características de forma son características que se basan en la forma del tumor. Para el presente trabajo se calcularon 8 características basadas en forma, como son: área, perímetro, longitud del eje mayor de la elipse, longitud del eje menor de la elipse, solidez, rectangularidad, delgadez y excentricidad [Kumar, 2017]; [Mufarroha, 2017]; [Sharif, 2018].

Región tumoral

La región tumoral se obtiene multiplicando la máscara binaria del tumor por la imagen original. A dicha región se le extrajeron las 4 características de intensidad, 14 de textura y 8 de forma, para un total de 26 características.

Aumento de la región tumoral

El aumento de la región tumoral utiliza la información proporcionada por los tejidos que rodean al tumor, ya que los tumores cerebrales de la misma categoría se alojan comúnmente en lugares similares. Se aplicó una dilatación morfológica con un elemento estructurante en forma de disco de radio 8 a la máscara binaria de la región tumoral; la imagen binaria resultante se multiplicó por la imagen original para obtener la región tumoral aumentada [Cheng, 2015], figura 1.



a) Región tumoral.

b) Aumento de la región tumoral.

Figura 1 Ilustración del aumento de la región tumoral, [Cheng, 2015].

A dicha región se le extrajeron las 4 características de intensidad y 14 de textura, para un total de 18 características. Debido a que la dilatación modifica el tamaño y la forma de la región tumoral original, se omitieron las características de forma.

Partición de la región tumoral aumentada

La partición de la región tumoral aumentada en subregiones en forma de anillo se utiliza para compensar la pérdida de información espacial. Consiste en calcular la distancia euclidiana más corta entre cada píxel y el borde de la máscara de la región aumentada; posteriormente las distancias de todos los píxeles se escalan linealmente a $[0, 1]$ y por último se divide el intervalo en 2^k subintervalos iguales para obtener 2^k máscaras en forma de anillo [Cheng, 2015]. Se particionó la región aumentada en 4 subregiones $\{A_0, A_1, A_2, A_3\}$ como se muestra en la figura 2, donde A_0 corresponde al anillo más grande y A_3 al más pequeño y la intersección de cualquier par de anillos es el conjunto vacío; de esta manera se evita la redundancia de datos. Cada máscara en forma de anillo fue multiplicada por la imagen original para extraerle las 4 características de intensidad y 14 características de textura, para hacer un total de 72 características.

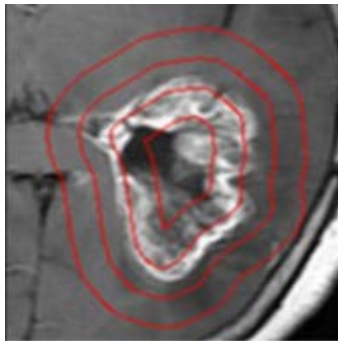


Figura 2 Partición en subregiones en forma de anillo, [Cheng, 2015].

Normalización

Dadas las diferentes características extraídas en el presente trabajo, es necesario realizar una normalización de datos, de esta manera se maneja una misma escala y se evita que la variación entre los rangos de las características interfiera en el resultado de la clasificación. Se decidió utilizar el método min-max, ya que tiene la ventaja de conservar todas las relaciones de los valores y no introducir ningún sesgo potencial a los datos [Liu, 2011]. Por lo tanto, todos los valores se escalan linealmente al intervalo $[0, 1]$.

Selección de características

Las características de una imagen suelen brindar gran cantidad de información, sin embargo, no todas tienen la misma importancia. Es por ello que uno de los principales objetivos de este trabajo es encontrar un subconjunto de características que proporcione una buena separabilidad de clases.

La selección de características es una de las técnicas frecuentes e importantes en el preprocesamiento de datos dentro del aprendizaje computacional; el proceso consiste en detectar las características más relevantes y eliminar datos irrelevantes, redundantes o ruidosos para mejorar la precisión predictiva [Kumar, 2014]. Debido que el número de combinaciones crece exponencialmente en función del número de características, se decidió utilizar tres enfoques para la selección, como son: selección mediante ranqueo utilizando el índice discriminante de Fisher; métodos de empaquetamiento como selección secuencial hacia adelante y hacia atrás; búsqueda heurística mediante un algoritmo genético.

El primer enfoque de selección de características consiste en puntuar y ordenar de forma descendente las características mediante el índice discriminante de Fisher. De este modo, se obtiene como primera propuesta de solución el subconjunto de características mejor puntuadas que maximizan la exactitud de clasificación.

Los métodos de empaquetamiento (wrapper) realizan una búsqueda en el espacio de posibles parámetros, dicha búsqueda requiere de un espacio de estado, un estado inicial, una condición de parada y un motor de búsqueda; su objetivo es encontrar el estado con la evaluación más alta, utilizando una función heurística para guiarlo [Kohavi, 1997]. Se utilizaron dos métodos wrapper de estrategia de búsqueda ávida como son: búsqueda secuencial hacia adelante (forward), donde se inicia con un conjunto vacío y progresivamente se agregan las características que maximizan la exactitud; búsqueda secuencial hacia atrás (backward), donde se inicia con el conjunto completo y progresivamente se eliminan las características menos prometedoras. De esta manera se obtienen otras dos propuestas de solución correspondientes a los subconjuntos obtenidos por ambos métodos wrapper.

La última propuesta de solución se obtiene mediante un algoritmo genético, los algoritmos genéticos son procesos de búsqueda utilizados en computación para

encontrar una solución exacta o aproximada a problemas de optimización combinatoria, ya que proporcionan una técnica para que el programa mejore automáticamente sus parámetros; están inspirados en la biología evolutiva, como la mutación hereditaria, la selección y el cruce [Kumar, 2010]. En el presente trabajo se empleó un algoritmo genético para seleccionar el subconjunto de características y los parámetros que maximizan la exactitud de clasificación para cada clasificador. Como se muestra en la figura 3, el cromosoma del individuo está representado mediante un arreglo binario; la parte izquierda corresponde a los parámetros del clasificador (en caso que requiera), la cual a su vez puede estar dividida en N secciones, donde N corresponde al número de parámetros del clasificador y cada parte representa un valor entero codificado en binario que corresponde al índice en un arreglo de valores; la parte derecha corresponde a las características seleccionadas, por lo tanto su longitud es igual al número de características totales, los 1's representan que la característica forma parte del subconjunto seleccionado y los 0's que no forma parte del subconjunto.

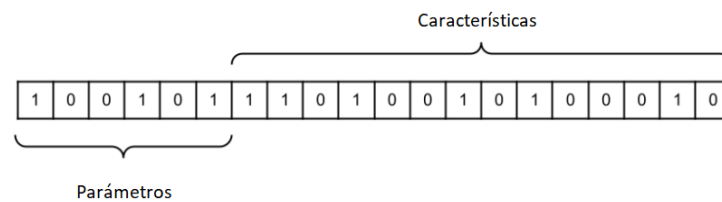


Figura 3 Cromosoma de un individuo de la población del algoritmo genético.

Clasificación

En el presente trabajo se busca obtener el mayor porcentaje de clasificación utilizando únicamente clasificadores tradicionales, para lograrlo se compara el desempeño de cinco diferentes clasificadores tradicionales: Bayesiano, Bayes ingenuo, k vecinos más cercanos (KNN), máquina de soporte vectorial (SVM) y bosques aleatorios (RF o random forest). Para cada clasificador se obtuvieron las cuatro propuestas de subconjuntos de características explicadas en el punto anterior para medir su rendimiento. El clasificador Bayesiano pertenece al conjunto de clasificadores estadísticos, se basa en el teorema de Bayes y predice la

probabilidad de que una muestra determinada pertenezca a una clase en particular [Kim, 2012].

El clasificador de Bayes ingenuo asume que un valor de atributo en una clase determinada es independiente a los valores de los otros atributos, por lo que se guía únicamente de las distribuciones estadísticas para calcular las probabilidades de verosimilitud [Rish, 2001].

La regla del vecino más cercano también conocido como clasificador KNN es uno de los algoritmos de aprendizaje supervisado más conocidos. Consiste en asignar al patrón desconocido la clase representada por la etiqueta mayoritaria de sus k vecinos más cercanos en el conjunto de entrenamiento [Gou, 2012]. La máquina de soporte vectorial se encarga de transformar los puntos de entrada a un espacio de características de una dimensión mayor; encuentra un hiperplano que los separa y maximiza el margen entre las clases en dicho espacio [Betancourt, 2005].

El clasificador Random forest consiste en una colección de clasificadores estructurados en árbol, donde cada clasificador contribuye con un solo voto para la asignación de la clase más frecuente [Rodríguez, 2012].

Para todos los clasificadores mencionados anteriormente, se utiliza la ampliamente conocida exactitud de clasificación como medida de rendimiento. Se escogió esta medida debido a que es la más utilizada en la literatura relacionada con la clasificación de tumores cerebrales usando el conjunto de datos analizado.

3. Resultados

La experimentación realizada en este trabajo sigue dos objetivos principales. El primero determina el subconjunto de características que aporta la información más relevante para realizar la clasificación. El segundo determina qué selector de características de los descritos en la sección 2 selecciona el mejor subconjunto de características y al mismo tiempo qué clasificador ofrece el mejor rendimiento.

Conjunto de imágenes

El conjunto público de imágenes cerebrales de resonancia magnética ponderadas en T1 se obtuvo del Hospital Nanfang, Guangzhou, China y del Hospital

General, Universidad Médica Tianjing, China, de 2005 a 2010 [Cheng, 2015]. El conjunto contiene 3,064 cortes pertenecientes a 233 pacientes, de los cuales 708 son meningiomas, 1,426 son gliomas y 930 corresponden a tumores pituitarios. Las imágenes tienen una resolución de 512 x 512 con un tamaño de píxel de 0.49 x 0.49 mm; el grosor y espacio de la rebanada es de 6 mm y 1 mm respectivamente. Se decidió utilizar dicho conjunto de imágenes debido a que contiene una suficiente cantidad de muestras, además brinda la segmentación realizada manualmente por el experto y proporciona cinco grupos estratificados para realizar la validación cruzada.

Efectividad de los conjuntos de características

Se experimentó utilizando tres diferentes conjuntos de características, para cada uno de estos se utilizó la ROI o máscara proporcionada por el experto, la cual permite segmentar y extraer características para el primer conjunto en particular. Para el segundo y tercer conjunto se utilizaron los métodos de aumento de la región tumoral y partición de la región tumoral aumentada, respectivamente (Sección 2). En la tabla 1 se muestra cada conjunto de características obtenido junto al número total de características extraídas.

Para el conjunto de características ROI se obtuvieron 26 características de las cuales, 4 son características de intensidad, 14 son características de textura y 8 características de forma (Sección 2).

El conjunto de características (ROI+) está conformada por 18 características: 4 de intensidad y 14 de textura. Finalmente, se combinaron las 26 características de la ROI con las 18 características obtenidas del aumento de la región, obteniendo un total de 44 características.

Tabla 1 Conjuntos de características obtenidos.

Conjunto de características	Número de características
Región de interés (ROI)	26
ROI y aumento de la región (ROI+)	44
ROI y anillos (ROI++)	98

Para el conjunto (ROI++), se realizó la partición de anillos con un total de 4 anillos, a cada anillo se le calcularon 18 características, siendo 4 características de intensidad y 14 características de textura para cada uno, además, se agregaron las 26 características de la ROI dando un total de 98 características. En la tabla 2 se muestran los resultados de exactitud de clasificación obtenidos por cada clasificador usando los diferentes conjuntos de características. Para estos resultados se experimentó, como primera aproximación, con el índice discriminante de Fisher. El número entre paréntesis de la tabla 2 indica el número de variables rankeadas por el FDR que maximizó el rendimiento.

Tabla 2 Resultados de exactitud de clasificación para cada conjunto de características.

Conjunto	Bayes	Bayes ingenuo	KNN	SVM	RF
ROI	77.69 (9)	75.15 (8)	79.81 (7)	88.52 (9)	80.13 (20)
ROI+	84.16 (10)	77.87(22)	87.74 (36)	89.18 (27)	86.90 (34)
ROI++	83.09 (13)	82.16 (10)	89.62 (51)	90.58 (40)	89.67 (40)

Los porcentajes de exactitud de clasificación más altos se encuentran resaltados en negrita en la tabla 2. Se observa que el conjunto ROI no contiene las mejores características para discriminar las clases del conjunto analizado. Sin embargo, para el conjunto de características de ROI+, se logra mejorar el rendimiento de todos los clasificadores respecto a lo obtenido con el conjunto ROI, donde destaca el mejor resultado para el clasificador Bayes con un total de 10 características, aunque dicho resultado no es competitivo con el resto de clasificadores restantes. En contraste, el conjunto ROI++ logra capturar las características más relevantes de los tumores analizados, ya que mejora el rendimiento de la mayoría de clasificadores, donde destaca la SVM que pasa de 88.52% a 90.58% utilizando 40 características.

Rendimiento de los métodos de selección de características

En este apartado se presentan los resultados obtenidos de los experimentos realizados con cada uno de los métodos de selección de características descritos en la sección 2. Se utilizó el conjunto de características ROI++ que obtuvo el mejor resultado para la mayoría de clasificadores en la tabla 2.

En la tabla 3 se muestran los porcentajes de exactitud de clasificación obtenidos por cada clasificador utilizando los diferentes métodos de selección de características. El número entre paréntesis indica el número de variables seleccionadas por el método correspondiente. Se puede observar que estas características son las que han capturado mejor la descripción de los tumores cerebrales analizados.

Tabla 3 Resultados usando diferentes métodos de selección de características.

Conjunto	Bayes	Bayes ingenuo	KNN	SVM	RF
FDR	83.09 (13)	82.16 (10)	89.62 (51)	90.58 (40)	89.67 (40)
Backward	86.22 (18)	86.43 (10)	90.43 (19)	91.01 (29)	89.39 (12)
Forward	88.73 (13)	86.30 (27)	90.64 (27)	91.13 (54)	88.73 (12)
Genético	89.14 (11)	87.71 (21)	91.27 (27)	92.35 (22)	90.75 (34)

De la tabla 3, se observa en general que al utilizar el algoritmo genético como selector de características y buscador de parámetros se obtienen los mejores resultados de exactitud de clasificación para todos los clasificadores evaluados. En particular, la SVM es la que mejor rendimiento ofrece (92.35%), el cual se alcanzó utilizando un kernel de base radial gaussiano y una tolerancia de 0.07, los parámetros *gamma* (0.036906) y constante C (1,686.714403) se obtuvieron por medio de una búsqueda por cuadrícula, el porcentaje de exactitud se obtuvo por medio de validación cruzada. Otro punto a destacar es el rendimiento competitivo que ofrece el método sencillo de KNN, lo cual indica la existencia de una relevancia alta en las características contenidas en el conjunto ROI++. Lo anterior remarca la importancia del uso de un algoritmo genético como selector de características.

Para analizar el rendimiento del clasificador SVM, se presenta la matriz de confusión en la tabla 4, donde se agrega también las medidas de sensibilidad y especificidad para mejor entendimiento de los resultados.

Tabla 4 Matriz de confusión para el clasificador SVM.

Clase real	Predicción			Sensibilidad
	1	2	3	
1	601	17	90	84.89
2	18	1,391	17	97.55
3	78	16	836	89.89
Especificidad	95.87	97.76	94.90	

Como se mencionó, en esta tabla se muestra la especificidad y sensibilidad de las tres clases de tumores: meningioma (clase 1), glioma (clase 2) y pituitario (clase 3) las cuales tienen una especificidad de 95.87, 97.76, y 94.90 y sensibilidad de 84.89, 97.55, y 89.89, respectivamente. Para cada clase, se observa que el principal problema de la clasificación está dado por las clases 1 y 3, ya que es entre estas donde el clasificador confunde más patrones. Lo anterior indica que existen algunas muestras de tumores de las clases 1 y 3 donde las características de forma, textura e intensidad extraídas del conjunto ROI++ son similares. En contraste, para la clase 2 se presenta un menor número de elementos clasificados erróneamente por lo que es en esta donde, tanto la sensibilidad como la especificidad, alcanzan sus valores más altos. Este resultado se considera adecuado e importante para el tipo de problema que se está analizando, ya que la clase 2 corresponde a tumores del tipo maligno, los cuales son los más peligrosos de entre las 3 clases y sobre los cuales se debería tener el mayor porcentaje posible de exactitud de clasificación y sensibilidad en particular, evitando problemas de falsas alarmas en cuanto a tumores malignos se refiere.

En la tabla 5 se muestran las características seleccionadas que mejoran la exactitud para el clasificador SVM, las cuales fueron encontradas mediante el algoritmo genético. Las características de la región tumoral y del anillo 0 (anillo más externo de la región aumentada) son las más relevantes debido a que se seleccionaron 10 y 7, respectivamente. De los anillos internos se seleccionaron únicamente 5 características, siendo el más pequeño el que no proporcionó ninguna. Esto confirma que la información de los tejidos circundantes al tumor brinda características discriminatorias complementarias a la región tumoral.

Tabla 5 Características seleccionadas para el clasificador SVM.

Región	Características seleccionadas
Región tumoral	Media, Asimetría, Correlación, Energía, Homogeneidad, Sombra de clúster, Área, Rectangularidad, Delgadez, Excentricidad.
Anillo 0	Asimetría, Correlación, Homogeneidad, Entropía, Momento de diferencia inversa, Suma promedio, Varianza de la suma.
Anillo 1	Inercia, Homogeneidad.
Anillo 2	Media, Entropía, Suma promedio.
Anillo 3	

4. Discusión

Los resultados obtenidos muestran que los clasificadores tradicionales proporcionan un buen porcentaje de clasificación al hacer uso de un algoritmo genético como selector de características usando los tejidos circundantes al tumor. En [Cheng, 2015], al usar el método de la partición de anillos se observa una exactitud de clasificación de 91.14% y 86.55% utilizando características sofisticadas y los clasificadores SVM y SRC (clasificación basada en representación dispersa) respectivamente, sin embargo en el presente trabajo se mejoran dichos resultados utilizando características sencillas de forma, textura e intensidad extraídas de la ROI y partición en anillos con un clasificador SVM y el algoritmo genético como selector de características, obteniendo un porcentaje de exactitud de clasificación del 92.35%.

5. Conclusiones

En el presente trabajo se realizó la clasificación de tres tipos de tumores cerebrales (meningioma, glioma y pituitario) comparando los resultados de distintos clasificadores como el Bayesiano, Bayes ingenuo, KNN, SVM y RF, los cuales utilizaron características de textura, intensidad y forma extraídas de la región tumoral, región tumoral aumentada y particionada en forma de anillos.

Los resultados obtenidos muestran que las características combinadas de la región tumoral con la región tumoral aumentada o región tumoral aumentada y particionada en anillos mejoran el porcentaje de exactitud para cada clasificador, resaltando entre ellos la máquina de soporte vectorial, la cual consigue el mayor porcentaje de exactitud (92.35%). Este resultado confirma que los tejidos que rodean al tumor brindan información adicional y relevante al momento de realizar la clasificación, ya que al tomar en cuenta subregiones pequeñas es posible obtener a detalle la información que brinda cada subregión aumentada del tumor. También, se observa que el método utilizado para la selección de características aumenta la influencia en la clasificación, logrando resultados muy diferentes al comparar los obtenidos mediante el FDR, los métodos wrapper y el uso del algoritmo genético, logrando los mejores resultados en este último.

Como trabajo futuro se considera extender el presente artículo en dos direcciones. La primera es comparar con más métodos de selección de características, así como utilizar métodos de ensamble de clasificadores. Otra alternativa es utilizar métodos de clasificación que incluyan la extracción de características de manera automática, como lo son las redes neuronales convolucionales.

6. Referencias

- [1] Abiwinanda N., Hanif M., Hesaputra S. T., Handayani A., & Mengko, T. R., (2018). Brain Tumor Classification Using Convolutional Neural Network. World Congress on Medical Physics and Biomedical Engineering, 2018, 183–189.
- [2] Afshar P., Mohammadi A., & Plataniotis K. N., (2018). Brain Tumor Type Classification via Capsule Networks. 2018 25th IEEE International Conference on Image Processing (ICIP).
- [3] Angulakshmi M. y Lakshmi Priya G., (2017). Automated brain tumour segmentation techniques - a review. International Journal of Imaging Systems and Technology, 27(1), 66-77.
- [4] Betancourt G. A., (2005). Las máquinas de soporte vectorial (SVMs). *Scientia et Technica*, 11(27), 67-72.
- [5] Cheng J., Huang W., Cao S., Yang R., Yang W., Yun Z., Wang Z., y Feng Q., (2015). Enhanced performance of brain tumor classification via tumor region augmentation and partition. *PLOS ONE*, 10(10), e0140381.
- [6] Gou J., Du L., Zhang Y., & Xiong T., (2012). A new distance-weighted k-nearest neighbor classifier. *Journal of Information & Computational Science*, 9(6), 1429-1436.
- [7] Haralick R. M., Shanmugam K., y Dinstein I., (1973). Textural features for image classification. *IEEE Transactions on Systems, Man, and Cybernetics*, SMC-3(6), 610-621.
- [8] Ismael M. R., y Abdel-Qader I., (2018). Brain tumor classification via statistical features and back-propagation neural network. In 2018 IEEE International Conference on Electro/Information Technology (EIT), 0252-0257.

- [9] Kim H. C., y Ghahramani Z., (2012). Bayesian classifier combination. In *Artificial Intelligence and Statistics*, pp. 619-627. PMLR.
- [10] Kohavi R., y John G. H., (1997). Wrappers for feature subset selection. *Artificial Intelligence*, 97(1-2), pp. 273-324.
- [11] Kumar M., Husain M., Upreti N., y Gupta D., (2010). Genetic algorithm: Review and application. *International Journal of Information Technology and Knowledge Management*, 2(2), pp. 451-454.
- [12] Kumar T. S., Rashmi K., Ramadoss S., Sandhya L. K., y Sangeetha T. J., (2017). Brain tumor detection using SVM classifier. In *2017 Third International Conference on Sensing, Signal Processing and Security (ICSSS)*, 318-323. IEEE.
- [13] Kumar V., y Minz S., (2014). Feature selection: a literature review. *SmartCR*, 4(3), 211-229.
- [14] Lala A., Gupta J. K., y Shringirishi M., (2013). Implementation on K-means Clustering and Fuzzy C-means Algorithm For Brain Tumor Segmentation. *International Journal of Computer Engineering & Science*, 3(1), 27-33.
- [15] Liu Z., (2011). A method of SVM with normalization in intrusion detection. *Procedia Environmental Sciences*, 11, 256-262.
- [16] Mufarroha F. A., Utamingrum F., & Mahmudy W. F., (2017). Segmentation Algorithm to Determine Group for Hand Gesture Recognition. *Journal of Telecommunication, Electronic and Computer Engineering (JTEC)*, 9(2-8), 13-17.
- [17] Qurat-UI-Ain G. L., Kazmi S. B., Jaffar M. A., y Mirza A. M., (2010). Classification and segmentation of brain tumor using texture analysis. *Recent advances in artificial intelligence, knowledge engineering and data bases*, 147-155.
- [18] Rish I., (2001). An empirical study of the naive Bayes classifier. In *IJCAI 2001 workshop on empirical methods in artificial intelligence*, Vol. 3, No. 22, pp. 41-46.
- [19] Rodriguez Galiano V. F., Ghimire B., Rogan J., Chica Olmo M., y Rigol Sanchez J. P., (2012). An assessment of the effectiveness of a random forest

- classifier for land-cover classification. *ISPRS Journal of Photogrammetry and Remote Sensing*, 67, 93-104.
- [20] Rehman A., Naz S., Razzak M. I., Akram F., y Imran M., (2020). A deep learning-based framework for automatic brain tumors classification using transfer learning. *Circuits, Systems, and Signal Processing*, 39(2), 757-775.
- [21] Saman S., y Narayanan S. J., (2019). Survey on brain tumor segmentation and feature extraction of MR images. *International Journal of Multimedia Information Retrieval*, 8(2), 79-99.
- [22] Sharif M., Tanvir U., Munir E. U., Khan M. A., y Yasmin M., (2018). Brain tumor segmentation and classification by improved binomial thresholding and multi-features selection. *Journal of Ambient Intelligence and Humanized Computing*, 1-20.
- [23] Zulpe N. y Pawar V., (2012). GLCM textural features for brain tumor classification. *International Journal of Computer Science Issues (IJCSI)*, 9(3), 354-359.