



Clasificación ordinal de los grados de afectación de la enfermedad de Parkinson empleando imágenes de transportadores presinápticos de dopamina

J. Camacho-Cañamón *
julio.camacho@uco.es

María-Victoria Guiote
UGC Medicina Nuclear §

Antonio-María Santos-Bueno
UGC Medicina Nuclear §

Ester Rodríguez-Cáceres
Equipo Provincial TICS §

Elvira Carmona-Asenjo
UGC Medicina Nuclear §

Juan-Antonio Vallejo-Casas
UGC Medicina Nuclear §
jantonio.vallejo.sspa@juntadeandalucia.es

P. A. Gutiérrez *
pagutierrez@uco.es

C. Hervás-Martínez *
chervas@uco.es

* *Dpto. de Informática y Análisis Numérico*
Universidad de Córdoba
Córdoba, España

§ *Hospital Universitario “Reina Sofía”*
Universidad de Córdoba. IMIBIC.
Córdoba, España

Resumen—La enfermedad de Parkinson se caracteriza por un descenso de la densidad de transportadores presinápticos de dopamina en los núcleos de la base. El método habitual de clasificación está basado en la observación y el análisis cualitativo de las imágenes obtenidas tras la administración de ^{123}I -ioflupano al paciente que se va a diagnosticar. De esta forma, las técnicas recientes de neuroimagen, como la imagen dopaminérgica utilizando tomografía computarizada por emisión de fotón único (SPECT-CT) con ^{123}I -ioflupano (DaTSCAN®), han demostrado detectar la enfermedad, incluso en etapas tempranas. Sin embargo, los comités internacionales recomiendan un análisis cuantitativo, asociado a la construcción de modelos de apoyo que complementen el diagnóstico visual. El objetivo del presente estudio es establecer un sistema de apoyo a la decisión, mediante la clasificación ordinal de las imágenes asociadas a los diferentes grados de afectación de la enfermedad mediante técnicas de aprendizaje automático. La base de datos utilizada está formada por 316 estudios realizados a pacientes entre septiembre de 2015 y mayo de 2018, distribuidos en tres grupos: 191 no padecen la enfermedad de Parkinson, 55 la padecen con un nivel de afectación leve y 70 con un nivel de afectación grave. Tras la administración intravenosa de 5 mCi (185 MBq), se realizó una SPECT-CT, preprocesando y normalizando espacialmente las imágenes. Como clasificador ordinal utilizamos un método de regresión logística, que nos permite obtener las características (vóxeles de la imagen) más informativas para la tarea de clasificación. El mejor modelo alcanzó un error absoluto medio máximo (MMAE) de 0,4857, tras la aplicación de un diseño experimental de tipo 5-fold. El análisis de los vóxeles más informativos, de acuerdo con el modelo obtenido, destaca regiones del cerebro que no son consideradas habitualmente por los especialistas para el

diagnóstico visual.

Index Terms—Enfermedad de Parkinson, SPECT-CT, aprendizaje automático, clasificación ordinal, regresión logística ordinal, imagen médica

I. INTRODUCCIÓN

Una de las características neuropatológicas de la enfermedad de Parkinson es un sustancial descenso de dopamina en los núcleos basales (caudado y putamen) debido a la disminución progresiva de la densidad de transportadores presinápticos [1]. La densidad de transportador presináptico de dopamina se puede detectar mediante técnicas de neuroimagen, que actualmente constituyen una práctica ordinaria en el diagnóstico de trastornos neurodegenerativos como la enfermedad de Parkinson. El ^{123}I -ioflupano (DaTSCAN®, General Electric Healthcare Limited, Little Chalfont. Bucks HP79NA U.K.) es un radiofármaco, ampliamente empleado para este fin, que se une a los transportadores presinápticos de dopamina en el cuerpo estriado y permite evaluar la densidad de estos con alta sensibilidad [2].

Actualmente, se combinan sistemas de cuantificación semi-automáticos, que analizan las imágenes, con los diagnósticos visuales dependientes de un observador especializado y, juntos, son capaces de distinguir entre las clases: control, patológico con nivel de afectación leve y patológico con nivel de afectación grave. En el caso del diagnóstico de la enfermedad de Parkinson, se aborda el problema con un enfoque clásico, es decir, mediante la cuantificación de la pérdida de dopamina neuronal en el estriado [3]. En este sentido, el Comité de la Asociación Europea de Medicina Nuclear y Neuroimagen recomienda el análisis cuantitativo, asociado a la construcción de modelos computacionales de apoyo que complementen el diagnóstico visual [4]. Debido a que la enfermedad de Parkinson afecta, fundamentalmente, a las neuronas que manejan

Este trabajo ha sido realizado gracias al apoyo económico derivado de los proyectos TIN2017-85887-C2-1-P y TIN2017-90567-REDT del Ministerio de Economía, Industria y Competitividad de España y, también, por el apoyo de los fondos FEDER de la Unión Europea. También se agradece a la Universidad de Córdoba por haber premiado al proyecto “Clasificación y evaluación automática de los grados de parkinson” UCO-SOCIAL-INNOVA, en el marco del cual se realiza este trabajo, en el III Plan Propio Galileo de Innovación y Transferencia. Los autores agradecen a NVIDIA Corporation la cesión de recursos computacionales a través del GPU Grant Program.

dopamina como neurotransmisor principal, la mayoría de los modelos analíticos asociados al uso de ioflupano se centran en el cuerpo estriado. En este estudio, nos centraremos en el cuerpo estriado, pero incluiremos el resto de la imagen cerebral en el análisis, dado que una hipótesis fundamental de este trabajo plantea la existencia de otras zonas del cerebro en las que los efectos de la dopamina también pueden ayudar a realizar una clasificación más precisa.

La mayoría de trabajos abordan el problema de clasificación binaria (pacientes de control y patológicos), utilizando imágenes funcionales, y empleando técnicas de aprendizaje automático basadas en regiones de interés (caudado y putamen) [5]. El análisis basado en estas regiones generalmente se justifica en que hay pérdida de actividad dopaminérgica en el putamen en relación con el caudado para los enfermos de Parkinson. Sin embargo, el rendimiento de modelos de clasificación binaria que utilizan solo estas regiones de interés es limitado. En primer lugar, no está claro si una variable aleatoria asociada a un vóxel (unidad volumétrica mínima de la imagen) es significativamente importante para la tarea de clasificación. Ciertos vóxeles, dentro de estas regiones, pueden sufrir una mayor pérdida de transportador de dopamina que otros y, por lo tanto, pueden ser más informativos. En segundo lugar, no se debe asumir que solamente el caudado y el putamen sean las zonas a partir de las cuáles se extraigan las características (vóxeles) que utilice el clasificador.

Se sabe que las estructuras extraestriadas están involucradas en la enfermedad de Parkinson. Y que, además, el globo pálido está involucrado en los subtipos de la enfermedad de Parkinson [6], así que esta región cerebral también se ha incluido en nuestro análisis. El hecho de incluir todos los vóxeles de estas regiones como características discriminantes reporta mayor interés que asumir *a priori* que ciertos vóxeles son los más importantes, según conocimiento especializado. Por lo tanto, nuestro estudio desarrolla un algoritmo basado en vóxeles en lugar de basado en las regiones de interés clásicas.

El modelo de aprendizaje automático utilizado para clasificación será una regresión logística ordinal regularizada [7]. Con la regularización del modelo se pretende minimizar el número de características que intervienen en la clasificación, es decir, vóxeles, mediante la minimización del valor absoluto de los coeficientes que tienen asociados. El modelo es una combinación lineal de los coeficientes asociados a los vóxeles, cuyo valor se emplea para calcular las probabilidades de pertenecer a las tres clases: normal (clase 1), patológico con nivel de afectación leve (clase 2) y patológico con nivel de afectación grave (clase 3). Así, mediante esta metodología estimaremos la probabilidad de pertenencia a una de las tres clases asociadas al nivel de afectación de la enfermedad de Parkinson. Además de la clasificación, otro de los objetivos de este estudio es reconocer cuáles son los “vóxeles informativos”, término que empleamos para denotar aquellos que son significativamente útiles para la clasificación, es decir, más determinantes para construir el clasificador. Estos serán un subconjunto de todos los vóxeles analizados de la imagen completa y, previsiblemente, también un subconjunto de todos

los vóxeles afectados por la enfermedad de Parkinson. Con el fin de crear modelos de clasificación aún más sencillos se evaluará el uso de un selector de características previo (ReliefF [8]).

Los resultados indican que el diagnóstico utilizando todos los vóxeles de la imagen es factible, obteniéndose un rendimiento aceptable para la complejidad del problema abordado. Además, se detectan vóxeles útiles para la tarea de clasificación en regiones que no habían sido previamente consideradas en la literatura. Por último, el uso del selector de características ReliefF no mejora el rendimiento del clasificador.

En la Sección II se realizará una revisión de los trabajos relacionados con el problema de clasificación de pacientes de parkinson. Posteriormente, en la Sección III, se detallarán los conjuntos de datos utilizados y su forma de obtención, así como la metodología empleada para generar los modelos de clasificación. Para concluir, en la Sección IV, se mostrarán los resultados obtenidos y se realizará una discusión de los mismos. La Sección V expondrá las conclusiones obtenidas tras este trabajo y las posibles futuras mejoras.

II. ESTADO DEL ARTE

Son muchos los métodos basados en aprendizaje automático utilizados en los sistemas de apoyo al diagnóstico médico, como pueden ser las máquinas de vectores soporte (MVS), la regresión logística (RL), los árboles de decisión, las redes bayesianas o las redes neuronales. Todas ellas son importantes en tanto en cuanto puedan ayudar al médico clínico a realizar un diagnóstico temprano de la enfermedad, permitan planificar su tratamiento y posibiliten la monitorización de la progresión de la enfermedad [5]. Estas metodologías están siendo ampliamente utilizadas en neuroimagen [9], debido a las siguientes ventajas: permiten la clasificación a nivel de individuo más que a nivel de grupo, por lo que los resultados tienen un alto potencial de traslación a la práctica clínica; son técnicas multivariantes y supervisadas que tienen en cuenta las características de los patrones (imágenes volumétricas, en este caso) distribuidas en un espacio de características complejo de alta dimensionalidad para entrenar el modelo de clasificación y, una vez entrenado, poder clasificar nuevos pacientes en función del modelo obtenido.

Algunos trabajos en los que se utilizan las MVS como clasificador binario [10], [11] están basados en extraer los vóxeles correspondientes al estriado. En [10], los autores realizan una descomposición de datos utilizando mínimos cuadrados parciales (*Partial Least Squares*) seguida por la utilización del clasificador MVS. En [11], los autores utilizan, una vez extraídos los vóxeles, un clasificador MVS lineal, contando con 208 imágenes DaTSCAN y utilizando máscaras para la selección de vóxeles.

Por otra parte, el modelo de RL determina la probabilidad de tener la enfermedad de Parkinson para un paciente, utilizando como características explicativas los vóxeles de las imágenes normalizadas. Esta metodología puede ser útil también para clasificar pacientes en diferentes categorías de riesgo de padecer la enfermedad, como sugieren algunos estudios utilizando



SPECT-CT en imágenes con ^{123}I -ioflupano, que representan la progresión de degeneración dopaminérgica de la enfermedad de Parkinson [12].

III. MATERIAL Y MÉTODOS

III-A. Material

Con el fin de evaluar la metodología propuesta en este trabajo, hemos utilizado los datos de 316 estudios con ^{123}I -ioflupano en pacientes derivados para evaluación de trastornos del movimiento realizados en el Hospital Universitario ‘Reina Sofía’ de Córdoba, en el periodo que va desde septiembre de 2015 hasta mayo de 2018. Los pacientes se distribuyen en tres clases: 191 no padecen la enfermedad de Parkinson (normal, clase 1), 55 padecen la enfermedad de Parkinson con un nivel de afectación leve (patológico, clase 2) y 70 padecen la enfermedad de Parkinson con un nivel de afectación grave (patológico, clase 3). La relación entre los valores muestrales asociados a padecer o no la enfermedad en cualquier de sus niveles de afectación se acercan a la relación poblacional entre dichas categorías. El diagnóstico definitivo se establece por la combinación de pruebas clínicas y complementarias realizadas por la Unidad de Neurología del centro hospitalario. La media de edad de los pacientes es de 70.46 años (35-89) con una desviación típica de 11,85 años. Las imágenes de SPECT-CT se adquieren según el protocolo indicado en [13].

Tres especialistas de la Unidad de Medicina Nuclear, han procesado, interpretado y evaluado las imágenes. La evaluación visual se ha establecido considerando exclusivamente el criterio normal y patológico en función del grado basándose en criterios comunes preestablecidos [14], y después de llegar a un informe de consenso entre los tres especialistas. Se consideró que un paciente era normal cuando, en la imagen volumétrica correspondiente a su estudio, aparecía una simetría bilateral en los núcleos basales, caudado y putamen; enfermedad de Parkinson con un nivel de afectación leve, cuando había una asimetría o reducción completa en la actividad del putamen; y con un nivel de afectación grave cuando había una ausencia bilateral de actividad en el caudado y putamen. Véase la figura 1, donde se muestra una imagen de un paciente normal en 1(a), la de un paciente patológico con un nivel de afectación leve en 1(b) y la de un paciente patológico con un nivel de afectación grave en 1(c).

III-B. Preprocesado de la imagen

El preprocesado de las imágenes consta de dos partes: *normalización espacial* y *recuperación de información*. Las imágenes suministradas inicialmente están sin procesar, por tanto hay que normalizarlas espacialmente, de forma que los mismos vóxeles representen las mismas características en todos los estudios. Para ello se utiliza el *software* PETRA [15] basado en el *framework* SPM8 [16], que permite leer y reunir varios archivos DICOM para formar una sola imagen espacial. Para realizar la normalización espacial se coordinan los ejes tridimensionales con la comisura anterior cerebral de la imagen, cuyas coordenadas están almacenadas en el

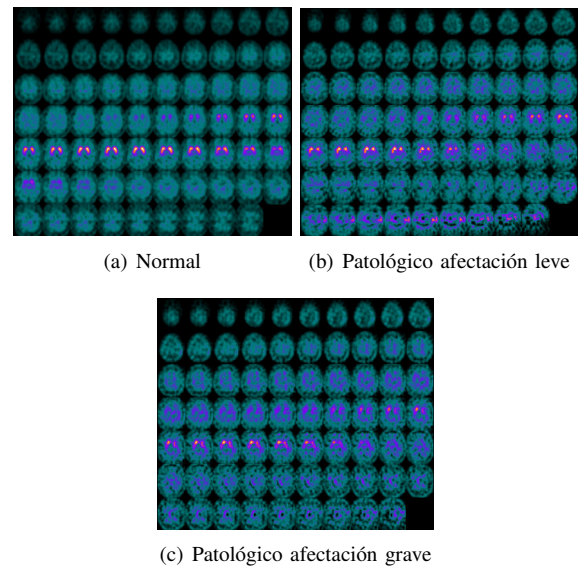


Figura 1. Imágenes de pacientes normal (a), patológico con nivel de afectación leve (b) y patológico con nivel de afectación grave (c).

formato DICOM de la imagen. No obstante, el método de normalización espacial se explica con detalle en [15].

Aplicando esta normalización a cada imagen, se obtiene una matriz tridimensional de $79 \times 95 \times 69$ vóxeles, es decir, $V = 517,845$ vóxeles por cada estudio. Este elevado número de vóxeles implica que es necesario aplicar técnicas de aprendizaje automático preparadas para un gran número de características. Tras la normalización espacial, es necesario hacer una búsqueda de valores perdidos en los vóxeles de cada imagen. Los valores perdidos serán sustituidos por la media de los valores que tienen los vóxeles del eje horizontal en el que se encuentra el valor perdido. A este proceso se le conoce como recuperación de información. Una vez preprocesadas las imágenes utilizaremos todos los vóxeles de cada paciente como características de entrada al clasificador, como se explica en la siguiente sección.

III-C. Metodología

III-C1. Notación y terminología: En primer lugar, definiremos algunos términos y notaciones. La imagen del paciente i -ésimo está representada por el vector \mathbf{x}_i , donde $i = 1, \dots, N$, siendo N el número de estudios (316). Si v es la posición de un vóxel concreto en una imagen, $v = 1, \dots, V$, entonces $\mathbf{x}_i(v)$ es el valor del vóxel v en la imagen del paciente i .

Leyendo la imagen de cada paciente i como una serie o conjunto de valores (vóxeles), la imagen \mathbf{x}_i se puede entender como un vector de tamaño $V \times 1$. Por tanto, $\mathbf{x}_i(v)$ es la componente v del vector de dimensiones $V \times 1$. Dicho de otra forma, cualquiera de estos vectores (de tamaño $V \times 1$) serviría para reconstruir una imagen $79 \times 95 \times 69$, la cual puede ser representada como una imagen 3D. De este modo, podremos considerar y visualizar un vector concreto.

Respecto a la clase de cada paciente i , la representaremos como una variable categórica y_i con tres categorías ordinales,

siendo $y_i = 1$ si el paciente i pertenece al grupo de control (normal), $y_i = 2$ si el paciente i padece enfermedad de Parkinson con un nivel de afectación leve, e $y_i = 3$ indicará que el nivel de afectación es grave.

Suponemos que este conjunto de patrones de entrenamiento es una realización de un conjunto de V variables aleatorias independientes e idénticamente distribuidas, así como de la variable asociada a la clase y_i de cada paciente. El conjunto de datos será representado por $D = \{(\mathbf{x}_i, y_i)\}, i = 1, \dots, N$. El número de pacientes de cada una de las clases de interés (normal, afección leve y afección grave) será denotado como N_1, N_2 y N_3 , respectivamente.

III-C2. Regresión logística ordinal regularizada: Consideraremos un modelo de RL ordinal [17] para estimar la probabilidad de que un paciente tenga un determinado nivel de afectación de la enfermedad de Parkinson, es decir, el modelo estima la probabilidad condicional $P(y_i = q|\mathbf{x}_i), q \in \{1, 2, 3\}$.

Suponemos que, al ser la variable dependiente ordinal, el mejor modelo que se ajusta a la nube de puntos es un modelo de RL ordinal. Como ya se mencionó, una aproximación similar, pero para clasificación binaria, [5] ha sido empleada para la detección de la enfermedad de Parkinson.

El modelo de RL ordinal (de forma más específica, *Proportional Odds Model*, POM, [17]) es un modelo de umbral [18] que modela la respuesta categórica ordinal mediante una función lineal de proyección $f(\mathbf{x}_i, \theta)$ común para todas las clases y un vector de umbrales β . Debido a que en nuestro problema hay tres clases, se cumple $\beta = \{b_1, b_2\}$, siendo $b_1 < b_2$. La probabilidad *a posteriori* de pertenencia a una clase se obtiene modelando la probabilidad de pertenencia acumulada a toda clase menor o igual a la evaluada:

$$P(y_i \leq q|\mathbf{x}_i, \hat{\mathbf{s}}) = \sum_{j=1}^q P(y_i = j|\mathbf{x}_i, \hat{\mathbf{s}}),$$

$$P(y_i = q|\mathbf{x}_i, \hat{\mathbf{s}}) = P(y_i \leq q|\mathbf{x}_i, \hat{\mathbf{s}}) - P(y_i \leq (q-1)|\mathbf{x}_i, \hat{\mathbf{s}}),$$

donde $\hat{\mathbf{s}} = \{\theta, \beta\}$ incluye todos los parámetros libres del modelo. Las probabilidades acumuladas se aproximan mediante la inversa de la función \logit (o función sigmoide):

$$P(y_i \leq q|\mathbf{x}_i, \hat{\mathbf{s}}) = \sigma(f(\mathbf{x}_i, \theta)) = \frac{1}{1 + \exp(-f(\mathbf{x}_i, \theta))},$$

para $q = \{1, 2\}$. Por definición, $P(y_i \leq 0|\mathbf{x}_i, \hat{\mathbf{s}}) = 0$ y $P(y_i \leq 3|\mathbf{x}_i, \hat{\mathbf{s}}) = 1$. La función de proyección es lineal, por lo que $f(\mathbf{x}_i, \theta) = \theta^T \cdot \mathbf{x}_i$. Por ello, la expresión final del modelo es:

$$\begin{aligned} P(y_i = 1|\mathbf{x}_i, \hat{\mathbf{s}}) &= \sigma(\theta^T \cdot \mathbf{x}_i - b_1), \\ P(y_i = 2|\mathbf{x}_i, \hat{\mathbf{s}}) &= \sigma(\theta^T \cdot \mathbf{x}_i - b_2) - \sigma(\theta^T \cdot \mathbf{x}_i - b_1), \\ P(y_i = 3|\mathbf{x}_i, \hat{\mathbf{s}}) &= 1 - \sigma(\theta^T \cdot \mathbf{x}_i - b_2). \end{aligned} \quad (1)$$

Para la estimación de los parámetros $\hat{\mathbf{s}} = \{\theta, \beta\}$, se utiliza el método de máxima verosimilitud, que minimiza la siguiente función de entropía cruzada:

$$L(\mathbf{s}, D) = -\frac{1}{N} \sum_{i=1}^N \sum_{q=1}^3 [[y_i = q]] \ln(P(y_i = q|\mathbf{x}_i)),$$

donde $[[c]] = 1$ si la condición c es cierta ($[[c]] = 0$, en caso contrario).

Es bien conocido que un sobre-ajuste del conjunto D puede incurrir en un modelo con varianza mayor, empobreciendo el rendimiento del conjunto de *test*. Una forma de controlar este fenómeno de sobre-entrenamiento es el uso de un término de regularización, que evita valores altos de los coeficientes, reduciendo la varianza y aumentando el sesgo, a costa de empeorar el ajuste al conjunto de entrenamiento. Para ello, modificamos la función de error de la siguiente forma:

$$L_2(\mathbf{s}, D) = L(\mathbf{s}, D) + \lambda \cdot \sum_{i=1}^V \theta_i^2, \quad (2)$$

donde λ es el parámetro de regularización. Dicho parámetro, que toma valores positivos, representa la importancia que se le da a la regularización frente a la minimización del error y debe ser ajustado por el investigador (normalmente en base a algún tipo de proceso de validación). Utilizando las Ecuaciones (1) y (2), el aprendizaje del modelo puede realizarse mediante las derivadas analíticas de $L_2(\mathbf{s}, D)$ con respecto a cada uno de los parámetros $s_i \in \{\theta, \beta\}$, mediante algún método de optimización (en el caso de la RL, se suelen emplear métodos de segundo orden). Estas derivadas pueden consultarse en [19]. Tal y como se discute en [20], la formulación de este modelo de RL ordinal se ajusta a una función de pérdida de tipo *Immediate-Threshold* (umbrales adyacentes), de forma que, para cada ejemplo etiquetado con y_i , solo se penaliza el error que cometen los umbrales que limitan el segmento correcto (es decir, b_{y_i} y b_{y_i-1}).

A partir de este modelo de RL, se pueden obtener, lo que hemos denominado, *Componentes Logísticas Principales* (CLP), que son aquellos vóxeles cuyo respectivo coeficiente θ_i es más elevado en valor absoluto. Es decir, estos vóxeles son los que más han influido en la función $f(\mathbf{x}_i, \theta)$, por lo que son los vóxeles más importantes para clasificar el sujeto en cualquiera de las tres clases consideradas. Las CLP tienen una estructura de componentes binarias de dimensión $V \times 1$, donde 1 significa que ese vóxel es relevante y 0 que no lo es. Por lo tanto, como hemos explicado previamente, se pueden llegar a representar como una imagen 3D. Así, se puede apreciar qué zonas son las que más afectan a la clasificación según el nivel de afectación de la enfermedad de Parkinson, como se ilustrará en el apartado de resultados.

III-C3. Diseño experimental: Para la evaluación de la bondad del modelo, aplicamos un diseño experimental de validación cruzada de tipo *5-fold*. Los pliegues o *folds* se crearán de forma estratificada, es decir, manteniendo la proporción original de patrones de pacientes normales, leves y graves.

Por cada una de las 5 particiones, realizaremos el aprendizaje estimando el mejor valor para el parámetro λ mediante una validación cruzada anidada: el conjunto de entrenamiento (que incluye 4 pliegues) se divide, de nuevo, en 2 conjuntos estratificados, aplicando un diseño de tipo *2-fold* por cada valor del parámetro λ que va a ser explorado. Los valores explorados para λ son los siguientes $\lambda \in \{10^{-3}, 10^{-2}, \dots, 10^3\}$.



A la hora de realizar el ajuste de λ elegimos el valor que produzca un menor error de validación. En concreto, utilizamos la función de error $MMAE$ (*Maximum Mean Absolute Error* [21], la cual es específicamente adecuada para problemas desequilibrados de clasificación ordinal). El MAE es una medida de error que tiene en cuenta la ordinalidad de la variable objetivo, $MAE = \frac{1}{N} \sum_{i=1}^N |y_i - y_i^*|$, donde $|y_i - y_i^*|$ es la distancia absoluta entre las etiquetas reales y predichas. El MAE varía desde 0 hasta 2 (que sería la máxima desviación en número de categorías en nuestro problema). Sin embargo, en problemas desequilibrados, las clases más frecuentes pueden dominar el error MAE , enmascarando un rendimiento pobre para las clases menos comunes. Es por ello que el $MMAE$ se define de la siguiente forma, $MMAE = \max\{MAE_1, MAE_2, MAE_3\}$, donde MAE_q es el error MAE teniendo en cuenta solo los patrones de la clase q : $MAE_q = \frac{1}{N_q} \sum_{i=1}^N |y_i - y_i^*|$, $q \in \{1, 2, 3\}$.

Con la intención de mostrar la mejora obtenida usando todos los vóxeles de las imágenes en lugar de usar solamente las regiones descritas por los especialistas, se han realizado pruebas con la misma experimentación, pero seleccionando previamente aquellas variables que los especialistas sitúan en las zonas estriales del cerebro.

Igualmente, se han aplicado técnicas de selección de características de tipo filtro, en concreto, el método ReliefF [8]. Estas son aplicadas como paso previo al entrenamiento del modelo con el fin de vislumbrar aquellas características con las que el modelo podría obtener mejores resultados.

Por último, una vez estimado el error mediante el proceso *5-fold*, estimaremos las CLP. Para ello, repetiremos el entrenamiento del modelo, pero considerando todos los pacientes disponibles, con el fin de obtener el mejor modelo posible. De nuevo, se aplicará una validación cruzada anidada de tipo *2-fold* que ajuste el mejor valor de λ en la RL ordinal. En nuestro experimento, para establecer las CLP, escogemos el 0,50% de los vóxeles (2590) cuyo θ_i es mayor en valor absoluto.

IV. RESULTADOS

Cada uno de los subconjuntos de entrenamiento generados en el procedimiento *5-fold* se utiliza para calcular la matriz de confusión de su respectivo conjunto de generalización. La suma de las matrices de confusión generadas para cada subconjunto de generalización será el resultado final del modelo, y a partir de ella se calculará el $MMAE$ completo. En la tabla I se muestran todos los resultados obtenidos empleando distintas metodologías: considerar todos los vóxeles (y dejar que la regularización seleccione los más importantes), considerar solo las regiones de interés determinadas por los expertos (caudado y putamen) y considerar el método ReliefF de selección de características manteniendo un 5%, 2% y 1% de los vóxeles originales. La tabla incluye una referencia a la subfigura de la Figura 2 que incluye la matriz de confusión correspondiente.

Los resultados demuestran que usando todos los vóxeles se consigue el mejor clasificador en $MMAE$ y en CCR . También cabe destacar que usando las regiones de interés que los especialistas han indicado, se consigue un $MMAE$ igual al que

Tabla I
ANÁLISIS DE RESULTADOS EN FUNCIÓN DE LOS VÓXELES CONSIDERADOS

Vóxeles	Figura	Nº Vóxeles	$MMAE \downarrow$	$CCR \uparrow$
Todos	2(a)	517845	0,4857	0,7532
Caudado y putamen	2(b)	3267	0,6727	0,7532
ReliefF (5%)	2(c)	25892	0,6000	0,7468
ReliefF (2%)	2(d)	10356	0,6727	0,7247
ReliefF (1%)	2(e)	5178	0,7091	0,7342

se consigue empleando ReliefF y seleccionando el 2% de las características más representativas. Sin embargo, el número de vóxeles empleados analizando las regiones de interés es una tercera parte de los que se utilizan con el selector. Esto indica que ReliefF no es adecuado para este problema, posiblemente por estar pensado para clasificación nominal.

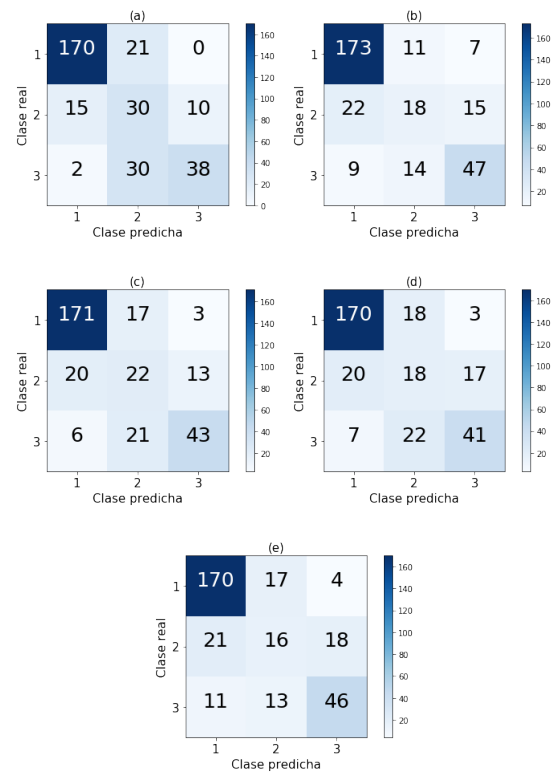


Figura 2. Matrices de confusión de los resultados de cada experimento.

Para finalizar, se visualizan los vóxeles más informativos de la imagen (CLP) tras entrenar los clasificadores con todos los datos y todos los vóxeles (ver Figura 3). Como puede apreciarse, existe una gran cantidad de vóxeles importantes agrupados en los propios núcleos basales, algo que era de esperar antes de realizar la representación. No obstante, debemos prestar especial atención a aquellas zonas que, sin pertenecer a las regiones de interés clásicas, también son muy relevantes para el modelo de RL ordinal en su tarea de clasificación. Estas lesiones extraestriales coinciden con la tendencia de otros estudios (resonancia magnética, por ejemplo) a evidenciar alteraciones en áreas cerebrales fuera

del estriado. Se corresponden a pequeñas zonas corticales, de predominio temporal y en la región media del lóbulo parietal.

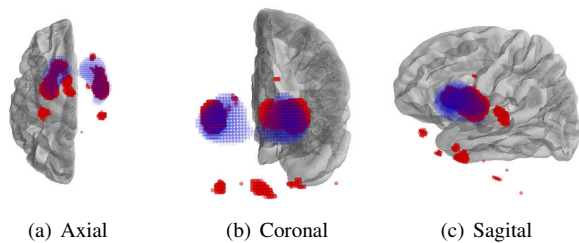


Figura 3. Representación 3D de los vóxeles más importantes (CLP), utilizando PySurfer [22]. En azul se representan los núcleos basales, caudado y putamen. En rojo se representan las CLP.

V. CONCLUSIÓN

Este artículo presenta un sistema de apoyo a la decisión mediante modelos de clasificación ordinal de pacientes de párkinson a partir de imágenes DaTSCAN. Para su desarrollo, se han empleado técnicas de aprendizaje automático, que incluye RL ordinal, y técnicas de normalización espacial y de recuperación de información. El modelo matemático computacional se considera más adecuado para el campo de la biomedicina dada su alta interpretabilidad.

Los resultados avalan que la metodología que utilizamos, empleando todos los vóxeles para la clasificación (junto con regularización en el modelo), obtiene un mejor rendimiento en la clasificación que si solamente se empleasen los vóxeles de las regiones de interés (núcleos basales) o si se emplease un selector de características nominal (ReliefF).

Respecto a la representación en 3D de las CLP, se confirma que los núcleos basales tienen una gran importancia en la tarea de clasificación de un paciente. Pero también se destacan otras zonas corticales que deberían de tenerse en cuenta para futuros estudios en la evaluación de la enfermedad de Parkinson. La densidad de transportadores presinápticos de dopamina tiene una gran importancia en la tarea de clasificación de un paciente, pero habría que tener en cuenta para futuros estudios otras zonas corticales que tendrían influencia en el diagnóstico diferencial de la enfermedad de Parkinson.

Como futuros trabajos se plantean realizar modificaciones en el método de selección de características nominal (ReliefF) para adaptarlo a una clasificación ordinal, de modo que la selección de características ayude a reducir la dimensionalidad del problema y mejorar los resultados. También se propone la aplicación de otros clasificadores como Procesos Gaussianos con núcleo lineal que permitan estimar todos los parámetros del modelo y calcular su incertidumbre.

REFERENCIAS

[1] R. Simon, D. Greenberg, and M. Aminoff, "Clinical neurology 5th edition," *Clinical Neurology*, 2009.

[2] J. Booij, J. B. Habraken, P. Bergmans, G. Tissingh *et al.*, "Imaging of dopamine transporters with iodine-123-fp-cit spect in healthy controls and patients with parkinson's disease," *The Journal of Nuclear Medicine*, vol. 39, no. 11, p. 1879, 1998.

[3] A. Antonini, K. L. Leenders, P. Vontobel, R. P. Maguire, J. Missimer, M. Psylla, and I. Günther, "Complementary pet studies of striatal neuronal function in the differential diagnosis between multiple system atrophy and parkinson's disease." *Brain: a journal of neurology*, vol. 120, no. 12, pp. 2187–2195, 1997.

[4] J. Darcourt, J. Booij, K. Tatsch, A. Varrone, T. Vander Borght, Ö. L. Kapucu, K. Nägren, F. Nobili, Z. Walker, and K. Van Laere, "Eann procedure guidelines for brain neurotransmission spect using 123i-labelled dopamine transporter ligands, version 2," *Eur J Nucl Med Mol*, vol. 37, no. 2, pp. 443–450, 2010.

[5] R. Prashanth, S. D. Roy, P. K. Mandal, and S. Ghosh, "Automatic classification and prediction models for early parkinson's disease diagnosis from spect imaging," *Expert Systems with Applications*, vol. 41, no. 7, pp. 3333–3342, 2014.

[6] A. Rajput, H. Sitte, A. Rajput, M. Fenton, C. Piffl, and O. Hornykiewicz, "Globus pallidus dopamine and parkinson motor subtypes clinical and brain biochemical correlation," *Neurology*, vol. 70, no. 16 Part 2, pp. 1403–1410, 2008.

[7] F. Pedregosa-Izquierdo, "Feature extraction and supervised learning on fmri: from practice to theory," Ph.D. dissertation, Université Pierre et Marie Curie-Paris VI, 2015.

[8] M. Robnik-Šikonja and I. Kononenko, "Theoretical and empirical analysis of relieff and rrelieff," *Machine learning*, vol. 53, no. 1-2, pp. 23–69, 2003.

[9] B. Magnin, L. Mesrob, S. Kinkingnéhun, M. Pélégriani-Issac, O. Colliot, M. Sarazin, B. Dubois, S. Lehéricy, and H. Benali, "Support vector machine-based classification of alzheimer's disease from whole-brain anatomical mri," *Neuroradiology*, vol. 51, no. 2, pp. 73–83, 2009.

[10] F. Segovia, J. Górriz, J. Ramírez, I. Alvarez, J. Jiménez-Hoyuela, and S. Ortega, "Improved parkinsonism diagnosis using a partial least squares based approach," *Medical physics*, vol. 39, no. 7, pp. 4395–4403, 2012.

[11] I. Illán, J. Górriz, J. Ramírez, F. Segovia, J. Jiménez-Hoyuela, and S. Ortega Lozano, "Automatic assistance to parkinson's disease diagnosis in datscan spect imaging," *Medical physics*, vol. 39, no. 10, pp. 5971–5980, 2012.

[12] A. Winogrodzka, P. Bergmans, J. Booij, E. Van Royen, A. Janssen, and E. C. Wolters, "[123 i] fp-cit spect is a useful method to monitor the rate of dopaminergic degeneration in early-stage parkinson's disease," *Journal of neural transmission*, vol. 108, no. 8, pp. 1011–1019, 2001.

[13] A. V. García, J. C. Vaamonde, V. G. Poblete, S. M. Rodado, M. R. Cortés, S. S. Ruiz, R. A. Ibáñez, and A. C. Soriano, "Utility of dopamine transporter imaging (123-i ioflupane spect) in the assessment of movement disorders," *Revista española de medicina nuclear*, vol. 23, no. 4, pp. 245–252, 2004.

[14] G. Perlaki, S. Szekeres, G. Orsi, L. Papp, B. Suha, S. A. Nagy, T. Doczi, J. Janszky, K. Zambo, and N. Kovacs, "Validation of an automated morphological mri-based 123i-fp-cit spect evaluation method," *Parkinsonism & related disorders*, vol. 29, pp. 24–29, 2016.

[15] F. Segovia, I. Á. Illán, D. Salas-Gonzalez, F. J. Martínez-Murcia, C. Phillips, C. G. Puntonet, J. R. P. de Inestrosa, and J. M. G. Sáez, "Petra: Multivariate analyses for neuroimaging data." in *IWBBIO*, 2014, pp. 1302–1312.

[16] W. D. Penny, K. J. Friston, J. T. Ashburner, S. J. Kiebel, and T. E. Nichols, *Statistical parametric mapping: the analysis of functional brain images*. Academic press, 2011.

[17] P. McCullagh, "Generalized linear models," *European Journal of Operational Research*, vol. 16, no. 3, pp. 285–292, 1984.

[18] P. A. Gutiérrez, M. Pérez-Ortiz, J. Sánchez-Monedero, F. Fernández-Navarro, and C. Hervás-Martínez, "Ordinal regression methods: survey and experimental study," *IEEE Transactions on Knowledge and Data Engineering*, vol. 28, no. 1, pp. 127–146, 2016.

[19] P. A. Gutiérrez, P. Tiño, and C. Hervás-Martínez, "Ordinal regression neural networks based on concentric hyperspheres," *Neural Networks*, vol. 59, pp. 51–60, 2014.

[20] J. D. Rennie and N. Srebro, "Loss functions for preference levels: Regression with discrete ordered labels," in *Proceedings of the IJCAI multidisciplinary workshop on advances in preference handling*. Kluwer Norwell, MA, 2005, pp. 180–186.

[21] M. Cruz-Ramírez, C. Hervás-Martínez, J. Sánchez-Monedero, and P. A. Gutiérrez, "Metrics to guide a multi-objective evolutionary algorithm for ordinal classification," *Neurocomputing*, vol. 135, pp. 21–31, 2014.

[22] Waskom, Gramfort, Burns, Luessi, and Larson, "Pysurfer," 2017. [Online]. Available: <http://pysurfer.github.io>