

## Trabajo Fin de Máster

Análisis de datos obtenidos por tomografía  
óptica de retina y su posible relación con la  
Esclerosis Múltiple y el Parkinson

Autor

Héctor Gracia Cabrera

Director/es

José Luis Pina Martínez

Francisco José Serón

Escuela de Ingeniería y Arquitectura

2013

# **Análisis de datos obtenidos por tomografía óptica de retina y su posible relación con la Esclerosis Múltiple y el Parkinson**

## **RESUMEN**

Este TFM se centra en evaluar las diferencias observadas en las diez capas que existen en la retina, comparando las de los pacientes con Esclerosis múltiple o Parkinson y los pacientes sanos. Para ello se emplea un nuevo prototipo de segmentación del dispositivo de Tomografía de Coherencia Óptica tipo Spectralis.

Los datos a analizar pertenecen a 463 pacientes, 204 con Esclerosis múltiple, 121 con Parkinson y 138 sanos. En total, cada paciente tiene asociados 7680 valores y un campo en el que se indica si es un paciente sano o que enfermedad padece.

Se desarrolló una herramienta que permite diferentes visualizaciones de los datos para poder observar con comodidad los datos de los diferentes pacientes, con el objetivo de poder validar estos datos. Este proceso de validación fue clave para alcanzar los resultados finales.

Se decidió probar a clasificar los pacientes usando un tipo concreto de red neuronal, los llamados support vector machine (SVM), por estar considerados como los mejores clasificadores binarios dentro de las redes neuronales.

Se utilizó un SVM para cada enfermedad y, tras varias pruebas, se consiguió que estos clasificadores obtuvieran precisiones superiores al 90% a la hora de diferenciar entre pacientes sanos y enfermos de ambas enfermedades.

El trabajo se completó con varias pruebas extra y se añadieron ciertos datos estadísticos que ofrecen más información sobre la validez del método desarrollado como criterio diagnóstico para estas enfermedades.

Además, se describen en los anexos las prácticas realizadas que son requisito para aprobar el Máster en Ingeniería Biomédica.

# Índice de contenido

<b>Capítulo 1. Introducción.....</b>	<b>3</b>
1.1 Oftalmología y Tomografía óptica.....	3
1.2 Enfermedades neurológicas consideradas.....	4
1.2.1 Esclerosis Múltiple (EM).....	4
1.2.2 Enfermedad de Parkinson (EP).....	4
1.3 Hipótesis del trabajo. “Posible relación entre tomografía óptica y enfermedades neurológicas”.....	5
1.4 Objetivo del TFM.....	5
<b>Capítulo 2. Datos, validación y propuesta de clasificación.....</b>	<b>8</b>
2.1 Datos existentes.....	8
2.2 Análisis visual de la información.....	9
2.2.1. Interfaz.....	11
2.3 Validación visual de los datos.....	13
2.4 Clasificación usando SVM.....	14
2.4.1. ROC.....	15
<b>Capítulo 3. Resultados obtenidos.....</b>	<b>17</b>
3.1 Experimentos realizados.....	17
3.2 Precisiones alcanzadas, y otros estadísticos.....	18
3.3 Area Under Curve (ROC).....	21
<b>Capítulo 4. Conclusiones.....</b>	<b>23</b>
<b>Capítulo 5. Trabajo futuro.....</b>	<b>24</b>
<b>Capítulo 6. Descripción de los anexos.....</b>	<b>25</b>
<b>Anexo A. Diagrama de Gant.....</b>	<b>26</b>
<b>Anexo B. Calculo de los datos estadísticos.....</b>	<b>27</b>
<b>Anexo C. Resultados de una clasificación.....</b>	<b>28</b>
<b>Anexo D. Prácticas realizadas.....</b>	<b>31</b>
<b>Bibliografía.....</b>	<b>36</b>

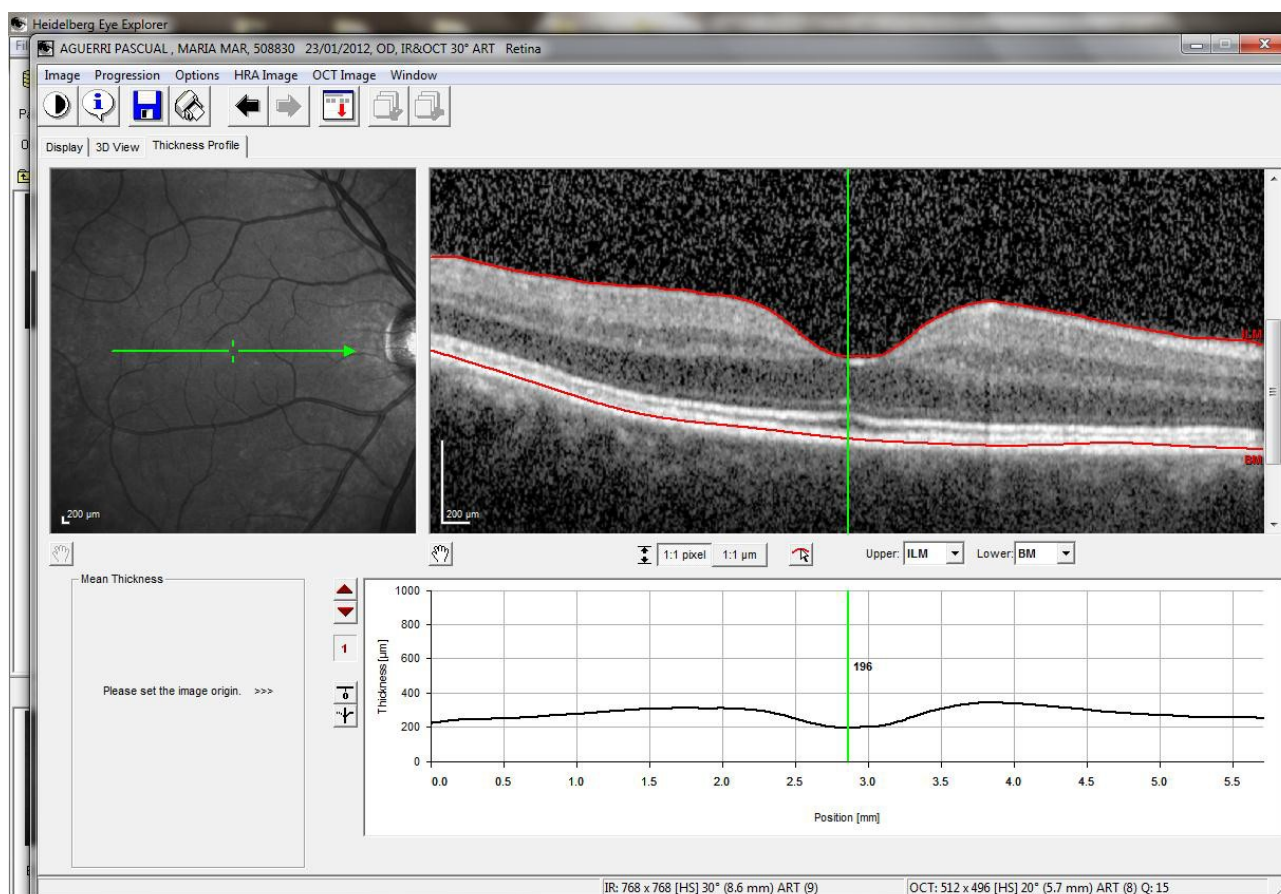
# Capítulo 1. Introducción.

Este capítulo define el contexto en el que se desarrolla el proyecto, tanto a nivel profesional como tecnológico, y especifica los objetivos del mismo.

## 1.1 Oftalmología y Tomografía óptica

La oftalmología es la especialidad médica que estudia las enfermedades del globo ocular, la musculatura ocular, sistema lagrimal, párpados y sus tratamientos. Esta especialidad ha tenido un gran impulso en las últimas décadas debido a la importancia atribuida al órgano de la visión y al espectacular avance tecnológico.

Un ejemplo de este avance tecnológico es el nuevo prototipo de segmentación del dispositivo de Tomografía de Coherencia Óptica tipo Spectralis. Este dispositivo utiliza una técnica de imagen tomográfica óptica, no invasiva e interferométrica.



*Fig. 1. Aspecto del terminal del Spectralis.*

Su precisión es tan elevada que se está estudiando la posibilidad de utilizarla para diagnosticar enfermedades neurológicas, al poder observar la degradación de las fibras nerviosas del ojo.

## 1.2 Enfermedades neurológicas consideradas

Como indica el título de este TFM, las enfermedades que se han tenido en cuenta han sido la Esclerosis Múltiple y el Parkinson, dos enfermedades cuyo origen y cura todavía son desconocidos.

### 1.2.1 Esclerosis Múltiple (EM)

La Esclerosis Múltiple<sup>[15]</sup> es una enfermedad neurológica que se suele presentar en adultos jóvenes.

La evolución de la enfermedad es muy variable: mientras en unos casos permite realizar una vida prácticamente normal, en otros la calidad de vida puede verse condicionada.

En España hay 40.000 personas con Esclerosis Múltiple, en Europa 500.000 y más de dos millones en el mundo.

La esclerosis múltiple (EM) es una enfermedad del sistema nervioso central (SNC) en el que se diferencian dos partes principales: cerebro y médula espinal. Envolviendo y protegiendo las fibras nerviosas del SNC hay un material compuesto por proteínas y grasas llamado mielina que facilita la conducción de los impulsos eléctricos entre las fibras nerviosas.

En la EM la mielina se pierde en múltiples áreas dejando, en ocasiones, cicatrices (esclerosis). Estas áreas lesionadas se conocen también con el nombre de placas de desmielinización.

Sólo puede ser diagnosticada con fiabilidad mediante una autopsia post-mortem o una biopsia, aunque existen criterios no invasivos para diagnosticarla con aceptable certeza. Son una serie de criterios destinados a decidir si se considera o no a una persona como afectada por la Esclerosis Múltiple, sin necesidad de recurrir a biopsias.

### 1.2.2 Enfermedad de Parkinson (EP)

La enfermedad de Parkinson<sup>[16]</sup> (EP) es una enfermedad neurodegenerativa del sistema nervioso central cuya principal característica es la muerte progresiva de neuronas en una parte del cerebro denominada sustancia negra pars compacta.

No se conoce la causa de esta enfermedad, la mayoría de los pacientes con enfermedad de Parkinson tienen una presentación esporádica.

En la actualidad se puede resumir que la EP no tiene una causa única, existiendo casos familiares con anomalías genéticas conocidas y otros casos en los que una conjunción de factores genéticos y ambientales serían los responsables de la muerte neuronal.

La EP se caracteriza por ser un trastorno motor progresivo cuyos principales síntomas son la torpeza generalizada con lentitud en la realización de movimientos, escasez de motilidad espontánea, temblor de reposo y rigidez.

En cualquier caso, debe hacerse énfasis en la enorme variabilidad en la evolución de la enfermedad de Parkinson, de forma que en algunos pacientes el proceso neurodegenerativo progresa muy lentamente mientras que en otros el proceso puede ser más rápido.

La edad de máxima presentación es a partir de los 60 años aumentando progresivamente hasta los 80 años. La incidencia entre estas edades es de un 1-2/1.000

de la población. Sin embargo, existen casos de presentación a edades más tempranas, pudiendo ocurrir a cualquier edad a partir de la segunda década. Afecta prácticamente por igual a ambos sexos.

En la actualidad, la EP es una enfermedad de diagnóstico clínico, puesto que no se ha identificado ningún marcador biológico de esta enfermedad. Esto significa que es el neurólogo quien, con los datos aportados por el paciente y su familia en la anamnesis y los hallazgos de la exploración física, diagnostica la existencia de una EP.

No es estrictamente necesaria la realización de pruebas complementarias (análisis de sangre, resonancia magnética o PET -estudio que permite conocer el estado de la función dopaminérgica cerebral-), si bien en determinados pacientes puede ayudar a aumentar la certeza diagnóstica y diferenciar la enfermedad de Parkinson de otros procesos que comparten características clínicas comunes.

La realización de un PET cerebral puede ayudar en aquellos pacientes con características atípicas en los que el diagnóstico resulta más difícil. En el momento actual no es una prueba de rutina ni disponible en todos los centros.

### **1.3 Hipótesis del trabajo. “Posible relación entre tomografía óptica y enfermedades neurológicas”**

El trabajo se centrará en evaluar las diferencias observadas en las diez capas que existen en la retina comparando las de los pacientes con Esclerosis múltiple o Parkinson y los pacientes sanos. Para ello se emplea un nuevo prototipo de segmentación del dispositivo de Tomografía de Coherencia Óptica tipo Spectralis.

### **1.4 Objetivo del TFM**

El objetivo de este TFM forma parte de los objetivos específicos de una investigación más amplia en la que está participando el grupo de investigación de prevención de la ceguera del servicio de oftalmología del Hospital Universitario Miguel Servet.

En la actualidad las principales líneas de este grupo se relacionan con el estudio estructural y funcional del nervio óptico en los siguientes aspectos:

- Evaluación transversal y longitudinal de la neuropatía óptica glaucomatosa.
- Diagnóstico y seguimiento de enfermedades neurodegenerativas mediante el estudio morfofuncional del nervio óptico.
- Neuroprotección y neuroregeneración mediante la administración de fármacos intraoculares. Penetración y Biodisponibilidad.
- Función visual relacionada con la cirugía oftálmica y neuropatías ópticas.

La repercusión en términos de salud de las técnicas en desarrollo permitiría tanto el diagnóstico, seguimiento y control terapéutico de enfermedades propias del nervio óptico como de enfermedades generales como la Esclerosis Múltiple o el Parkinson. Todo ello mediante métodos rápidos, eficientes y totalmente incruentos.

El constante contacto y cooperación con la industria supone la implicación del grupo en el desarrollo y mejora de diversos sistemas diagnósticos y procedimientos de liberación de fármacos. Lo que ha desembocado en la firma de acuerdos de cooperación con las empresas líderes en los sectores implicados.

Como se puede comprobar en las publicaciones del grupo<sup>[2][3][4][5][6]</sup> la cooperación con grupos internacionales es extensa. Como más reciente evento en este capítulo cabe destacar la creación de un consorcio europeo para la solicitud de un ambicioso proyecto FP7 sobre *Detection, diagnosis and monitoring of different optic neuropathies and glaucoma*, junto a:

- UCL Institute of Ophthalmology. Prof David F. Garway-Heath
- University of Thessaloniki, Greece (Assist Prof Fotis Topouzis; Dr Panayiota Founti)
- University of Mainz, Germany (Dr Franz Grus)
- University of Aberdeen (Prof Augusto Azuara-Blanco)
- University of Genoa, Italy (Prof Michele Iester)
- City University London, UK (Prof David Crabb)
- Bietti Foundation, Rome (Dr Francesco Oddone)
- University Eye Hospital, Wurtzburg, Germany (Dr Marc Schargus)
- University Complutense Madrid, Spain (Prof Julián García-Feijoo)

Este TFM se inscribe en el actual trabajo en neurooftalmología de este grupo, cuyo objetivo general es evaluar la situación visual de los pacientes afectos de Esclerosis múltiple, Parkinson y Enfermedad de Alzheimer en distintos estadios de evolución y la utilidad de las mediciones de la capa de fibras nerviosas de la retina como marcador de diagnóstico, evolución y respuesta terapéutica.

Los objetivos específicos del trabajo de este grupo son:

1. Cuantificar la disfunción visual y neurológica en todos los pacientes con Esclerosis múltiple, Parkinson y Enfermedad de Alzheimer incluidos en este estudio.
2. Evaluar el daño axonal presente a nivel de la capa de fibras nerviosas de la retina en los distintos estadios de evolución de las enfermedades.
3. Evaluar la utilidad de las diferentes técnicas de análisis funcional y estructural del nervio óptico en la detección de daño secundario a dichas patologías y el grado de correlación existente entre ellas.

4. Analizar la capacidad del estudio estructural y funcional de la capa de fibras nerviosas de la retina como biomarcador de diagnóstico precoz y de evolución de la enfermedad.
5. Evaluar la correlación existente entre las mediciones de la capa de fibras nerviosas de la retina y los test de examen cognitivo y de disfunción funcional en los pacientes incluidos en el estudio.
6. Evaluar las modificaciones observadas en la capa de fibras nerviosas de la retina a lo largo del tiempo, en función de las diferentes actitudes terapéuticas y de la progresión de las enfermedades.
7. Evaluar las modificaciones observadas en las diez capas que existen en la retina de los pacientes con Esclerosis múltiple, Parkinson y Enfermedad de Alzheimer, comparados con los sujetos sanos, empleando para ello un nuevo prototipo de segmentación del dispositivo de Tomografía de Coherencia Óptica tipo Spectralis. Se pretende evaluar cada capa retiniana por separado para demostrar cuáles son las capas que se modifican en función de la neurodegeneración causada por cada una de las enfermedades a lo largo del tiempo y de las diferentes actitudes terapéuticas.

Es el último de estos objetivos el que se ha abordado en este TFM, excluyendo la Enfermedad de Alzheimer. La metodología seguida y los resultados se describirán a continuación.



## Capítulo 2. Datos, validación y propuesta de clasificación

En este capítulo se describirá el conjunto de datos disponible, como fueron validados y una posible solución que permite diferenciar los pacientes sanos de los enfermos basándose en estos datos.

### 2.1 Datos existentes

Los datos de la retina extraídos mediante una tomografía óptica coherente (OCT) consisten en los espesores asociados a cada una de las 10 capas que tiene la retina en 786 puntos, se encuentran en una base de datos con formato SPSS (programa estadístico de IBM).

La estructura de la retina ha sido ampliamente estudiada y se conoce a la perfección. Las 10 capas en las que normalmente se divide son:

1. Membrana limitante interna
2. Capa de fibras nerviosas de la retina
3. Células ganglionares
4. Capa plexiforme interna
5. Capa nuclear interna
6. Capa plexiforme externa
7. Capa nuclear externa
8. Membrana limitante externa
9. Fotorreceptores
10. Epitelio pigmentario retina

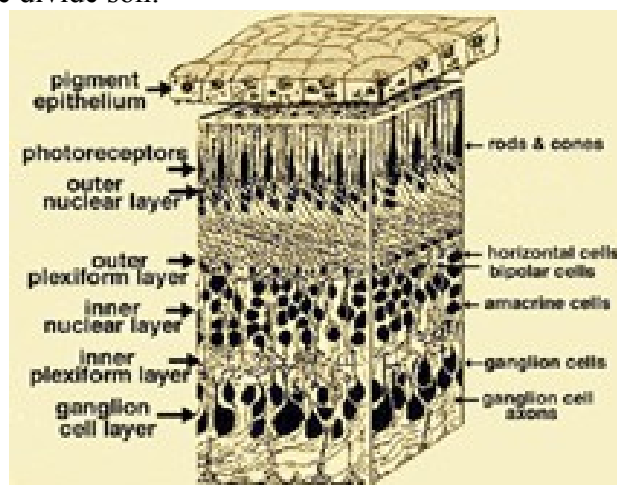


Fig. 2. Histología de la retina

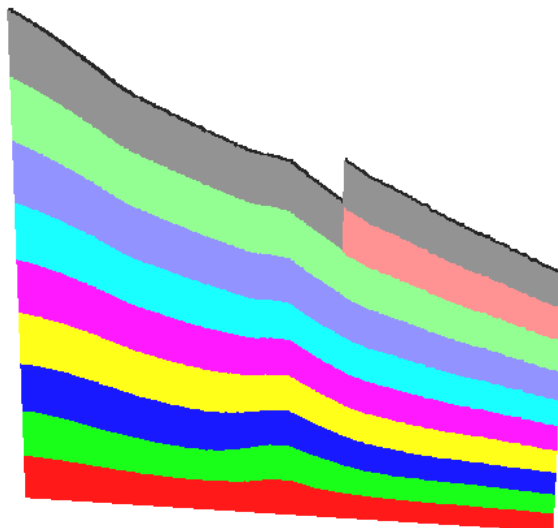
En total, cada paciente tiene asociados 7680 valores y un campo en el que se indica si es un paciente sano o que enfermedad padece (EM o EP). Los datos pertenecen a 463 pacientes, 204 con EM, 121 con EP y 138 sanos.

Dada la numerosa cantidad de información asociada a cada paciente, resulta imposible hacerse una idea de su estructura o distribución. Por ello, el primer objetivo del trabajo fue realizar una representación visual de esta información usando la herramienta Visualization Toolkit (VTK).

La Visualization Toolkit<sup>[13]</sup> (VTK) es un software de código abierto disponible gratuitamente para gráficos 3D, procesamiento de imágenes y su visualización. Este software permite múltiples lenguajes, de entre ellos se escogió Tcl<sup>[18][19]</sup> para desarrollar la herramienta de visualización.

## 2.2 Análisis visual de la información

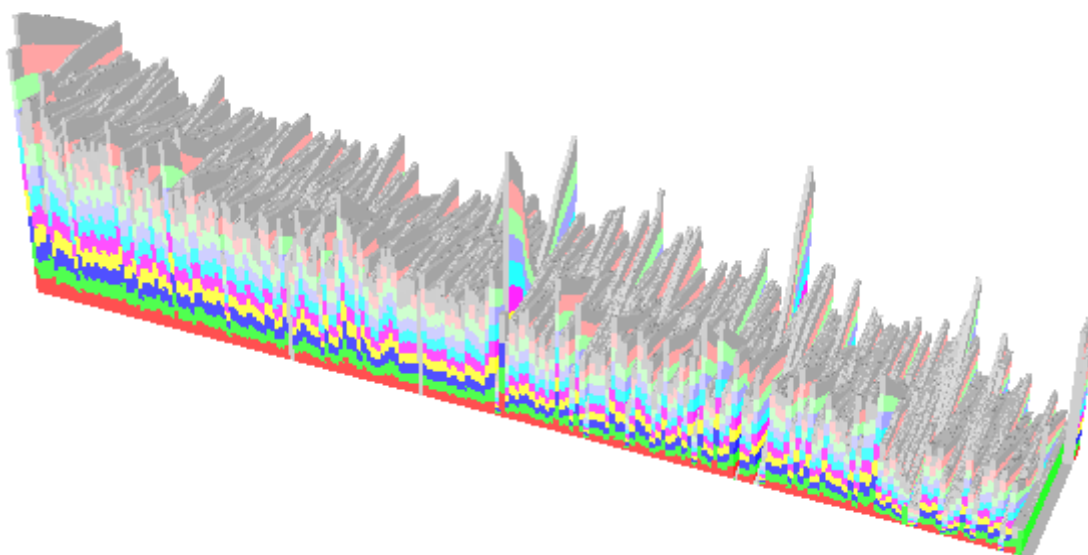
La primera meta a alcanzar fue representar la estructura de capas de un paciente. En el proceso se observó que, en la mayoría de los pacientes, solo los primeros 512 valores de cada capa contenían valores no nulos. Después de consultar con el equipo de médicos, se decidió utilizar únicamente estos primeros 512 valores y descartar el resto, consiguiendo mayor homogeneidad en los datos.



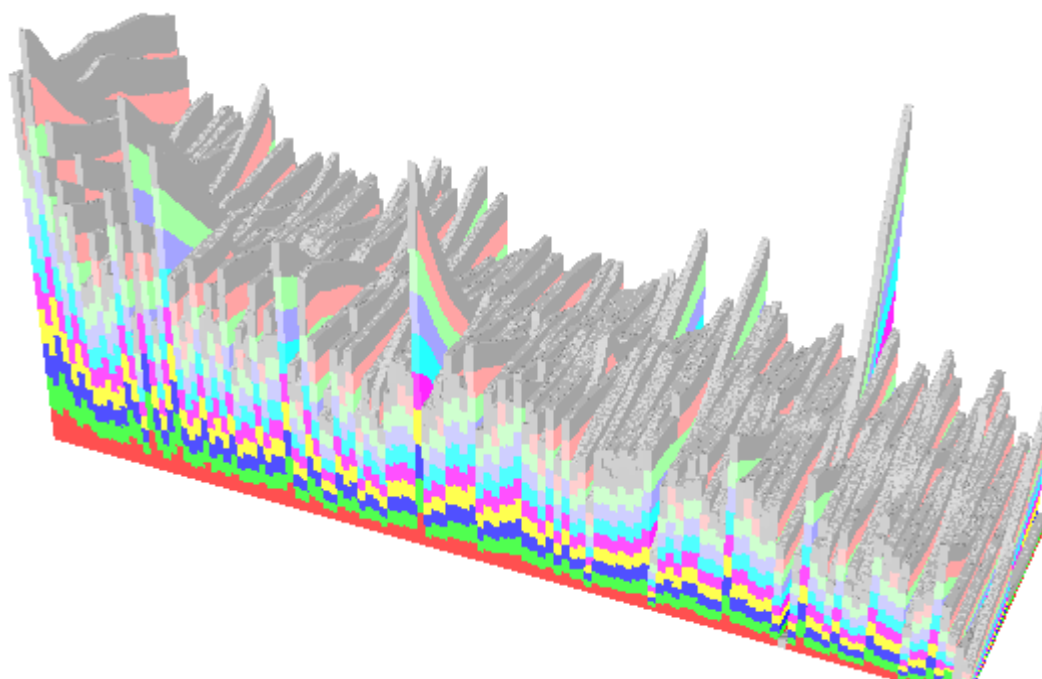
*Fig. 3. Estructura de capas de un paciente con 512 puntos por capa*

En la figura 3 pueden apreciarse los datos de un paciente, las capas se encuentran ordenadas de abajo a arriba. Puede observarse como la capa 9, de color anaranjado, contiene errores, ya que faltan los datos de la parte izquierda. Esta capa es la más difícil de delimitar y muchos pacientes tienen errores en ella. La máquina que extrae los datos todavía está en fase de prototipo y se espera que en su versión final pueda evitar o minimizar estos errores.

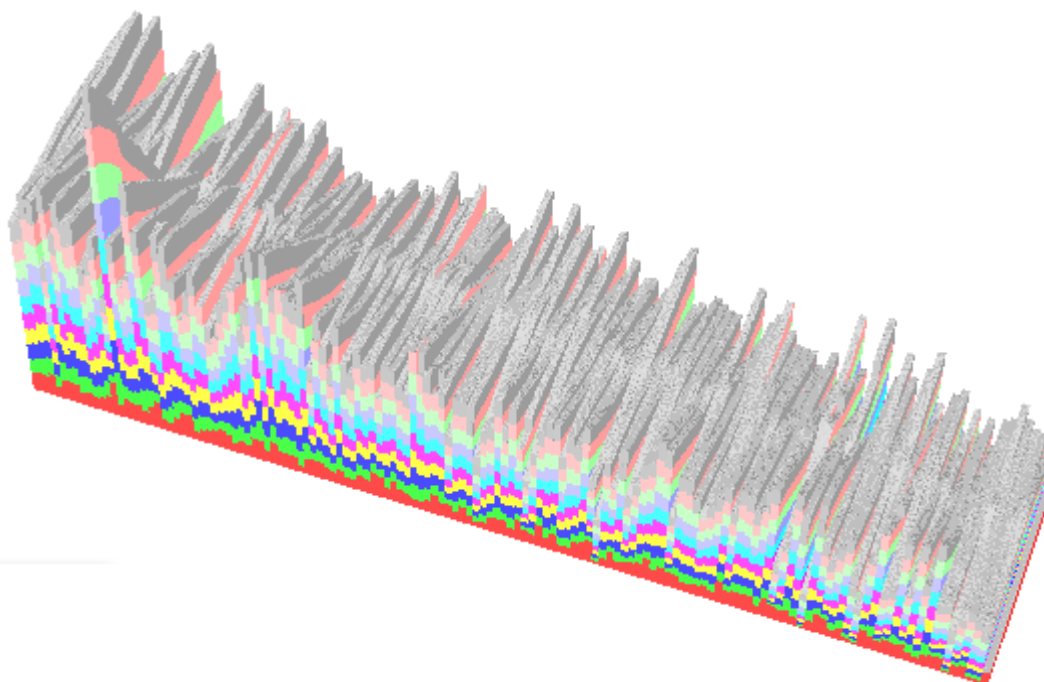
El siguiente paso fue representar un conjunto de pacientes, para comprobar si existía alguna relación entre los datos de cada uno de ellos. Para mejorar su visualización, se ordenaron en función de su altura media y se separaron en los tres conjuntos, EM, EP y sanos.



*Fig. 4. Enfermos de EM*



*Fig. 5. Enfermos de EP*



*Fig 6. Pacientes sanos*

En la versión final, se decidió representar las capas en orden inverso, la 10 abajo y la 1 arriba, ya que los médicos preferían verlo de ese modo, por eso los colores aparecerán en orden inverso en las siguientes figuras.

Esta herramienta se fue mejorando para permitir diferentes visualizaciones de los datos y se añadió una pequeña interfaz para poder visualizar con comodidad los diferentes pacientes. Esta interfaz se describirá brevemente a continuación.

### **2.2.1. Interfaz**

A continuación se describirán brevemente las funcionalidades de cada uno de los botones de la interfaz y posteriormente se mostrara una imagen de la versión final de la herramienta.

Es una interfaz bastante sencilla, pero suficiente para cumplir con su cometido, poder visualizar y validar los datos de una forma cómoda. Como podrá verse más adelante, fue una pieza clave que ayudo a mejorar los resultados del trabajo.

La herramienta es interactiva y permite mover los objetos en un entorno de tres dimensiones, además de las funciones descritas a continuación.

Reiniciar vista	- Deshace los cambios de las cámaras.	
Ver todos los pacientes	- Muestra todos los pacientes juntos.	
Paciente 57		
Agrupar 1	- Permite visualizar los diferentes pacientes uno a uno.	
Siguiente	- Muestra varios pacientes agrupados en forma de matriz.	
Anterior	- Muestra el siguiente paciente.	
Ver todas las capas	- Muestra el paciente anterior.	
<input checked="" type="checkbox"/> Capa 1	} Muestra/oculta las distintas .s.	
<input checked="" type="checkbox"/> Capa 2		
<input checked="" type="checkbox"/> Capa 3		
<input checked="" type="checkbox"/> Capa 4		
<input checked="" type="checkbox"/> Capa 5		
<input checked="" type="checkbox"/> Capa 6	} Bloquea las vistas seleccionadas para que no bien su contenido al usar los controles anteriores.	
<input checked="" type="checkbox"/> Capa 7		
<input checked="" type="checkbox"/> Capa 8		
<input checked="" type="checkbox"/> Capa 9		
<input checked="" type="checkbox"/> Capa 10	} Identificadores de los entes.	
<input type="checkbox"/> Bloquear vista de EM		
<input type="checkbox"/> Bloquear vista de EP		
<input type="checkbox"/> Bloquear vista de CEM		
<input type="checkbox"/> Bloquear vista de CEP		
Identificador del paciente con Esclerosis Multiple: 200		} Leyenda que indica el orden de las capas y su nombre.
Identificador del paciente con Enfermedad de Parkinson: 1087		
Identificador del paciente de Control de Esclerosis Multiple: 2103		
Identificador del paciente de Control de Enfermedad de Parkinson: 2120		
Capa 1- Inner glial limiting membrane		
Capa 2- Nerve fiber layer		
Capa 3- Ganglion cell layer		
Capa 4- Inner plexiform layer		
Capa 5- Inner nuclear layer		
Capa 6- Outer plexiform layer		
Capa 7- Outer nuclear layer		
Capa 8- Outer glial limiting membrane		
Capa 9- Photoreceptors		
Capa 10- Retinal pigment epithelium		

Fig. 7. Interfaz de la herramienta de visualización diseñada.

Como puede apreciarse en la siguiente imagen, aparecen 4 grupos de pacientes en vez de 3. Esto se debe a que se decidió separar los pacientes sanos en dos grupos, cada uno relacionado con una de las enfermedades que se están tratando. Esta separación se realizó en base a la edad de los pacientes sanos, ya que estas dos enfermedades las sufren personas de rangos de edad diferentes y bien diferenciados. En los resultados se podrá apreciar que esta fue una buena decisión.

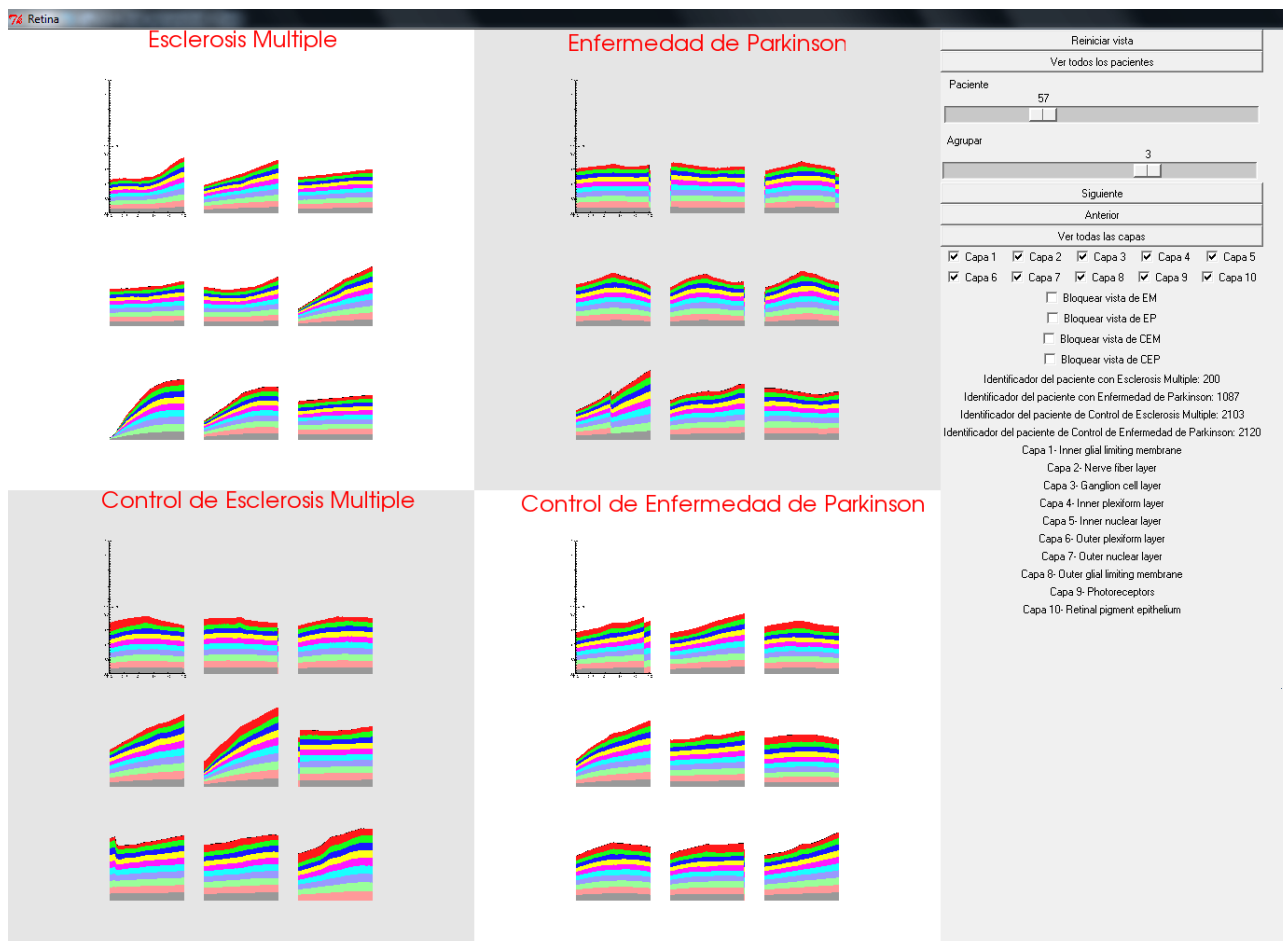


Fig. 8. Versión final de la herramienta, mostrando varios pacientes de cada tipo a la vez.

## 2.3 Validación visual de los datos

Gracias a esta herramienta, se comprobó la gran variabilidad de los datos, probablemente debida a la diversidad biológica. Además, se observó que muchos pacientes tenían zonas más gruesas en alguno de los extremos y aparentemente no seguían ningún patrón, podían localizarse a la izquierda, a la derecha o no presentarse.

Después de consultar con el equipo de médicos, se aclaró que esto era debido a que los datos se tomaban en diferente orden, según que ojo del paciente se estuviera analizando. Se decidió representar los ojos izquierdos como si fueran ojos derechos cambiando el orden de sus puntos. De este modo, las partes anatómicas representadas por cada punto, serán las mismas independientemente del ojo al que pertenezcan los datos. Para poder realizar esto, se visualizaron uno a uno todos los pacientes y se identificó a que ojo pertenecían los datos, con ello se añadió un nuevo campo a la base de datos donde se indicaba si el ojo era izquierdo o derecho. A la vez, se aprovechó para eliminar aquellos pacientes que tenían demasiados errores. En total se eliminaron 13 pacientes de los 463.

En este punto del trabajo se apreciaba con mayor claridad la dimensión del problema que pretendíamos resolver. Esperábamos que la visualización de los datos mostrara algún patrón o ciertas similitudes en los enfermos, que estableciera un punto de partida en la búsqueda de una solución que permitiera clasificarlos. Pero no se apreciaron grandes diferencias entre los datos.

Se decidió calcular y representar los valores medios de cada uno de los grupos,

para comprobar si mostraban las diferencias que estábamos buscando, pero los datos medios de cada grupo resultaron ser prácticamente idénticos, como se aprecia a continuación.

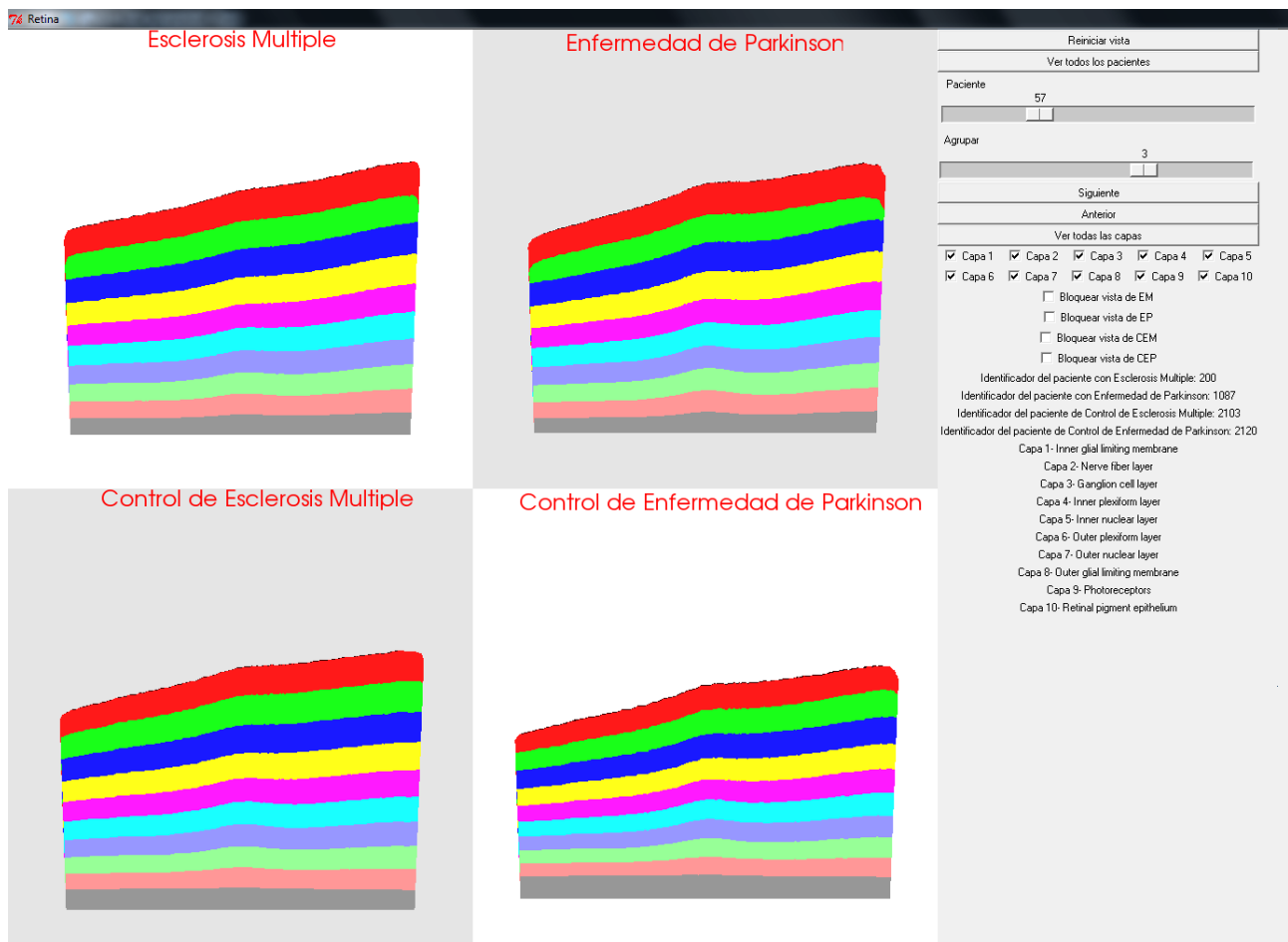


Fig. 9. Valores medios de los diferentes grupos de pacientes.

Ya que a simple vista no se podía establecer una regla para poder clasificar los pacientes, se decidió buscar algún método estadístico que permitiera tal clasificación. En algunos de los trabajos anteriores, el equipo de investigación médico, había utilizado redes neuronales como una posible herramienta diagnóstica<sup>[2][6]</sup>. Siguiendo su ejemplo, se decidió probar a clasificar los datos usando un tipo concreto de red neuronal, los llamados support vector machine (SVM), por estar considerados como los mejores clasificadores binarios dentro de las redes neuronales.

## 2.4 Clasificación usando SVM

Para realizar la clasificación de los datos se utilizaron dos Support Vector Machine<sup>[1][7][20]</sup> (SVM), uno para cada enfermedad. Concretamente se usó la versión para Matlab<sup>[14]</sup>, libsvm-mat.

Los SVM están reconocidos como los mejores clasificadores porque siempre encuentran la mejor división posible entre las dos clases de datos, enfermos y sanos en este caso, en cambio otros tipos de redes neuronales pueden finalizar en mínimos locales y no encontrar la división óptima.

Sin embargo los SVM, también tienen sus desventajas. Estos clasificadores tienen varios parámetros que permiten modificar sus características para que se ajusten

mejor a diferentes distribuciones de datos. Por lo tanto, optimizar estos parámetros será crucial para que se realice una buena clasificación de los datos.

Por suerte, se pueden optimizar de forma automática algunos parámetros a partir de los datos, pero otros pueden requerir una búsqueda manual en la que no existe ningún criterio formal para encontrar el óptimo.

Además de estos parámetros, también influyen en la clasificación otros factores como, la separación de los datos en datos para entrenar y datos para validar, el rango de valores de los datos o la correlación de la información de las diferentes variables.

Los SVM muestran como resultado la precisión alcanzada en la clasificación del grupo de validación, por ejemplo Accuracy = 96.875% (62/64) (classification), donde se aprecia que únicamente dos pacientes se han clasificado mal en este ejemplo. Sin embargo, a la hora de diagnosticar enfermedades, suelen usarse otros datos estadísticos que añaden información extra sobre la calidad del método diagnóstico. Uno de los más usados es el área under ROC curve (AUC), explicado brevemente a continuación.

### 2.4.1. ROC

La curva ROC<sup>[9][10]</sup> es una representación gráfica de la sensibilidad frente a (1 – especificidad) para un sistema clasificador binario. Otra interpretación de este gráfico es la representación del número de verdaderos positivos (VPR = Razón de Verdaderos Positivos) frente al número de falsos positivos (FPR = Razón de Falsos Positivos).

Los valores de sensibilidad y especificidad que representan las curvas ROC, por si solos también son un buen indicativo de la calidad de un método diagnóstico, ya que la sensibilidad es la probabilidad de clasificar correctamente a un individuo enfermo y la especificidad la probabilidad de clasificar correctamente a un individuo sano.

Para la elección entre dos pruebas diagnósticas distintas, se recurre a las curvas ROC, ya que es una medida global. Por esto, en el ámbito sanitario, las curvas ROC también se denominan curvas de rendimiento diagnóstico.

La elección se realiza mediante la comparación del área bajo la curva (AUC) de ambas pruebas. Este área posee un valor comprendido entre 0 y 1, donde 1 representa un valor diagnóstico perfecto y valores menores que 0,5 una prueba sin capacidad discriminativa diagnóstica. Por esto, siempre se elige la prueba diagnóstica que presente un mayor área bajo la curva (AUC).

A modo de guía para interpretar las curvas ROC se han establecido los siguientes intervalos para los valores de AUC, aunque estos valores pueden variar levemente según la fuente:

*Tabla 1. Calidad del test diagnóstico en función del AUC<sup>[25]</sup>*

AUC	[0.5, 6)	[0.6, 0.75)	[0.75, 0.9)	[0.9, 0.97)	[0.97, 1)
Calidad del test	Malo	Regular	Bueno	Muy bueno	Excelente

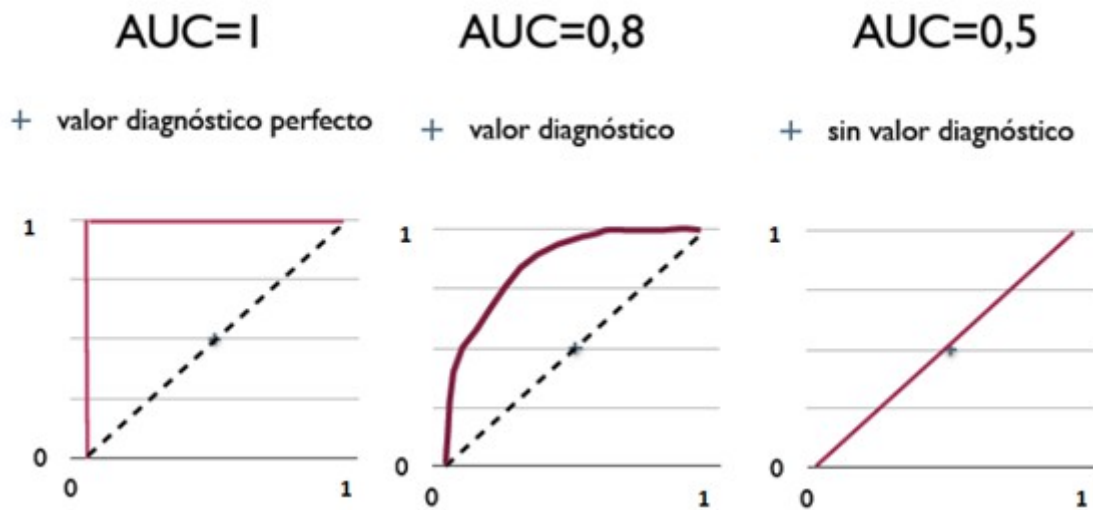
*Tabla 2. Calidad del test diagnóstico en función del AUC<sup>[11]</sup>*

AUC	[0.6, 0.7)	[0.7, 0.8)	[0.8, 0.9)	[0.9, 1)
Calidad del test	Malo	Sin valor	Bueno	Excelente

Además, su representación gráfica es muy explícita y fácilmente interpretable a



simple vista. A modo de ejemplo se muestran los casos extremos de este tipo de curvas.



*Fig.10. Ejemplos de curvas ROC*

# Capítulo 3. Resultados obtenidos

En este capítulo se explicaran los métodos utilizados para solucionar este problema y sus resultados.

## 3.1 Experimentos realizados

El método seguido para comprobar el objetivo de este TFM fue intentar clasificar los pacientes de ambas enfermedades usando un SVM para cada una de ellas, como se ha descrito anteriormente. Si se consigue una buena clasificación usando este método, se comprobará que efectivamente existe relación entre los datos extraídos por tomografía óptica y las enfermedades a estudio.

Siguiendo las especificaciones de los SVM, se crearon dos matrices de datos para cada enfermedad. Una contiene los valores de la base de datos y la otra las etiquetas que identifican a los pacientes como enfermos con 1 o sanos con -1. Por lo tanto se obtienen cuatro matrices, dos relacionadas con Esclerosis Múltiple y otras dos con la Enfermedad de Parkinson.

Una vez los datos están preparados, el siguiente paso consiste en separar los pacientes en dos conjuntos. Uno sera utilizado para entrenar los SVM y el otro para calcular su precisión predictora. Antes de separar los pacientes en estos dos conjuntos, se ordenaron de forma aleatoria para evitar que su distribución inicial afectara de algún modo a los resultados.

Uno de los conjuntos se utiliza para entrenar el SVM, indicándole a que clase pertenece cada uno de los pacientes con los que se entrena, con el objetivo de que pueda encontrar características que diferencien las dos clases.

Con los SVM ya entrenados, se utiliza el otro conjunto de pacientes para comprobar su calidad predictora. Sin indicarle a que clase pertenecen los pacientes, se introducen sus datos y el SVM los clasifica basándose en la información que extrajo del conjunto de entrenamiento. Como ultimo paso, se comprueba la cantidad de aciertos resultantes de esta clasificación.

Como primera prueba se utilizo la base de datos original, en la que no se habían eliminado pacientes, ni se tenia en cuenta a que ojo pertenecían los datos. Las precisiones alcanzadas en esta primera prueba fueron del 23% en EM y del 60% en Parkinson.

Posteriormente se realizó una prueba utilizando unicamente los primeros 512 puntos de cada capa, por los motivos explicados en el capítulo 2.2. Además, se modificaron los datos de ojos izquierdos para cada capa, como ya se ha descrito anteriormente y se realizó un escalado de los datos para que se encontraran en el intervalo  $[0,1]$  para mejorar la clasificación como indica la guía de los SVM<sup>[7]</sup>.

Esta modificación de los datos puede realizarse directamente en Matlab o con cualquier otro lenguaje de programación. En este caso, primero se exporto la base de

datos, que se encontraba en formato SPSS, a fichero de texto separado por tabulaciones y se modificaron los datos en ese fichero.

Para que las pruebas fueran más fluidas, se realizaron a la vez en varios SVM que estaban configurados con diferentes parámetros, con el objetivo de encontrar la mejor configuración posible. El parámetro principal de los SVM es el tipo de kernel o núcleo, que define el tipo de función que será utilizada para separar las diferentes clases. Se probaron tres configuraciones diferentes de este parámetro:

- Linear kernel: utiliza una función lineal.
- Precomputed kernel: utiliza una función definida por el usuario partiendo de los datos.
- RBF: utiliza una función radial similar a curvas gaussianas.

Además, se añadió una cuarta configuración llamada n-fold cross validation o validación cruzada<sup>[22][23]</sup>, esta configuración utiliza todos los datos para entrenar y validar, ordenados de n formas diferentes. Luego se obtiene una media de las n clasificaciones diferentes. Se suele usar la validación cruzada cuando el número de elementos a clasificar es pequeño, para comprobar si la separación de los datos en los conjuntos para entrenar y validar influye en los resultados y para obtener un valor de la precisión más realista. Altos valores de precisión en los resultados de la validación cruzada, suelen indicar que el modelo generaliza bien y probablemente se obtendrán buenos resultados con datos de nuevos.

Resumiendo, el proceso realizado se puede resumir en los siguientes pasos:

1. Se realizó un escalado de los datos para que sus valores se encuentren en el intervalo  $[0,1]$ .
2. Se aleatorizo el orden de los pacientes.
3. Los pacientes se separaron en dos conjuntos, uno para entrenar el SVM y otro para validarlo.
4. Se entrenan los SVM con uno de los conjuntos de pacientes.
5. Se comprueba la precisión alcanzada clasificando el otro conjunto de pacientes.
6. Se repite el proceso variando los parámetros del SVM o modificando los datos con el objetivo de maximizar la precisión alcanzada.

## 3.2 Precisiones alcanzadas, y otros estadísticos

A continuación se muestran varias tablas con las precisiones alcanzadas por las diferentes configuraciones y los pasos que llevaron a la configuración final.

Se utilizó el 80% de los pacientes para entrenar y el resto para validar en tres casos diferentes. Con los datos originales de las 10 capas, con los datos de las 10 capas escalados al rango de valores  $[0,1]$  y, por último, con los datos de 9 capas escalados al rango de valores  $[0,1]$ . En el último caso, se eliminó la capa 9 por contener demasiados errores en muchos pacientes.

Tabla 3. Precisiones de los diferentes SVM.

	10 capas		Datos entre [0,1], 10 capas		Datos entre [0,1], 9 capas	
	EM	EP	EM	EP	EM	EP
<b>Lineal</b>	82.608696	80.000000	80.000000	87.272727	92.173913	94.545455
<b>Precomputed</b>	82.608696	80.000000	80.000000	87.272727	92.173913	94.545455
<b>RBF</b>	75.652174	56.363636	75.652174	61.818182	89.565217	96.363636
<b>10-fold</b>	88.627451	88.205128	90.196078	89.743590	89.411765	89.230769

En el último de los casos se obtuvieron las mejores precisiones, además se aprecia que el kernel lineal obtiene mejores resultados en casi todos los casos y los resultados de la validación cruzada son muy elevados en todos los casos, lo que da un valor extra a los resultados obtenidos.

Después, se comprobó si la separación de los pacientes afectaba a la clasificación. Estas pruebas se realizaron sobre la última de las configuraciones anteriores. En la tabla se indica la cantidad de pacientes que se usaron para entrenar en cada caso, el resto se utilizó para validar.

Tabla 4. Variación de la precisión del SVM lineal.

	70% para entrenar		75% para entrenar		90% para entrenar	
	EM	EP	EM	EP	EM	EP
<b>Lineal</b>	93.506494	91.525424	95.312500	93.877551	92.307692	90.000000

La variación en la cantidad de datos usados para entrenar y validar no afectó en exceso a la precisión de los modelos, lo que implica que el modelo generalizaba bien y probablemente se obtendrán buenos resultados con datos de nuevos pacientes. Los resultados de la separación del 75% serán los que se utilizarán para calcular el resto de resultados que se muestran a continuación, por haber obtenido la mejor precisión.

En las siguientes tablas se muestran varios datos estadísticos calculados a partir de los resultados obtenidos con el SVM. Estos datos son el área bajo la curva ROC (AUC), el intervalo de confianza (CI) al 95% del valor del AUC, la sensibilidad, la especificidad, Likelihood Ratio positivo y negativo.

El 95% CI<sup>[8][11][12]</sup> asegura estadísticamente que el AUC se encontrará en ese rango de valores para el 95% de la población. Teóricamente, contra menor sea este intervalo, mejor funcionará con un grupo diferente de pacientes.

Los ratios likelihood<sup>[24]</sup> son test estadísticos usados para comparar el ajuste de dos modelos, el positivo se considera mejor contra más se aleja por encima de la unidad y el negativo mejor contra más se acerca a cero.

El resto de los datos estadísticos ya han sido descritos en la sección 2.4.1.

Los valores de la primera fila pertenecen a la clasificación realizada con el método explicado anteriormente. Los valores de la segunda fila corresponden a una clasificación realizada usando un único valor, el average thickness, es decir, el grosor medio de la capa de fibras nerviosas del ojo. Este valor era el que se intentaba utilizar como método diagnóstico de estas enfermedades, antes de que la tecnología fuera capaz de extraer los grosores de cada capa. Como puede observarse en los valores de AUC del segundo método, apenas proporciona información diagnóstica, al menos usándolo en un

SVM.

*Tabla 5. Datos estadísticos obtenidos para EM.*

Parámetros OCT	AUC	95% CI	Sensibilidad (%)	Especificidad (%)	+Likelihood Ratio	- Likelihood Ratio
9 capas, 512 puntos	0.921	0.889 – 0.952	100	83.33	5.99	0.166
Average thickness.	0.539	0.507 – 0.570	100	0	1	1

*Tabla 6. Datos estadísticos obtenidos para EP.*

Parámetros OCT	AUC	95% CI	Sensibilidad (%)	Especificidad (%)	+Likelihood Ratio	- Likelihood Ratio
9 capas, 512 puntos	0.928	0.892 – 0.963	100	88.24	8.50	0.117
Average thickness.	0.429	0.393 – 0.464	100	0	1	1

Para completar los resultados, se hicieron pruebas de clasificación utilizando los valores de cada capa por separado, para comprobar si existían capas más significativas que otras.

*Tabla 7. Precisiones alcanzadas con los valores de cada capa por separado para EM*

	10	9	8	7	6	5	4	3	2	1
<b>Lineal</b>	87.500	71.875	73.437	77.419	73.770	76.666	76.271	74.137	78.947	83.928
<b>RBF</b>	90.625	71.875	73.437	77.419	73.770	76.666	76.271	74.137	78.947	82.142
<b>10-fold</b>	84.313	76.862	77.254	76.862	76.862	76.862	76.862	76.862	76.862	77.254

En el caso de la EM las capas 1 y 10 parecen ser más significativas que las otras.

*Tabla 8. Precisiones alcanzadas con los valores de cada capa por separado para EP*

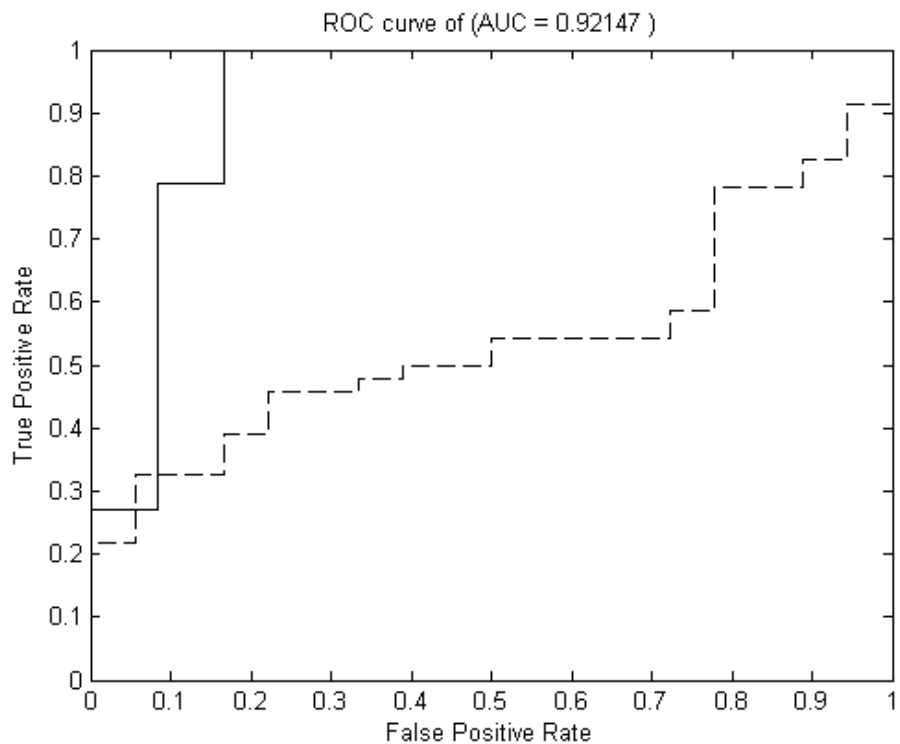
	10	9	8	7	6	5	4	3	2	1
<b>Lineal</b>	79.591	81.632	59.183	61.702	63.043	60.000	52.272	65.116	64.285	68.292
<b>RBF</b>	81.632	83.673	71.428	61.702	63.043	60.000	52.272	65.116	64.285	68.292
<b>10-fold</b>	81.538	78.974	61.538	60.000	60.000	60.512	58.461	60.512	60.000	61.02

En el caso de la EP las capas 9 y 10 parecen ser mucho más significativas que las otras.

Aun así, en ambas enfermedades se obtienen mejores resultados en la clasificación usando los valores de todas las capas, en vez de únicamente las más significativas. Sin embargo, esta información puede ser útil a la hora de comprobar que capas se ven más afectadas por cada una de las enfermedades.

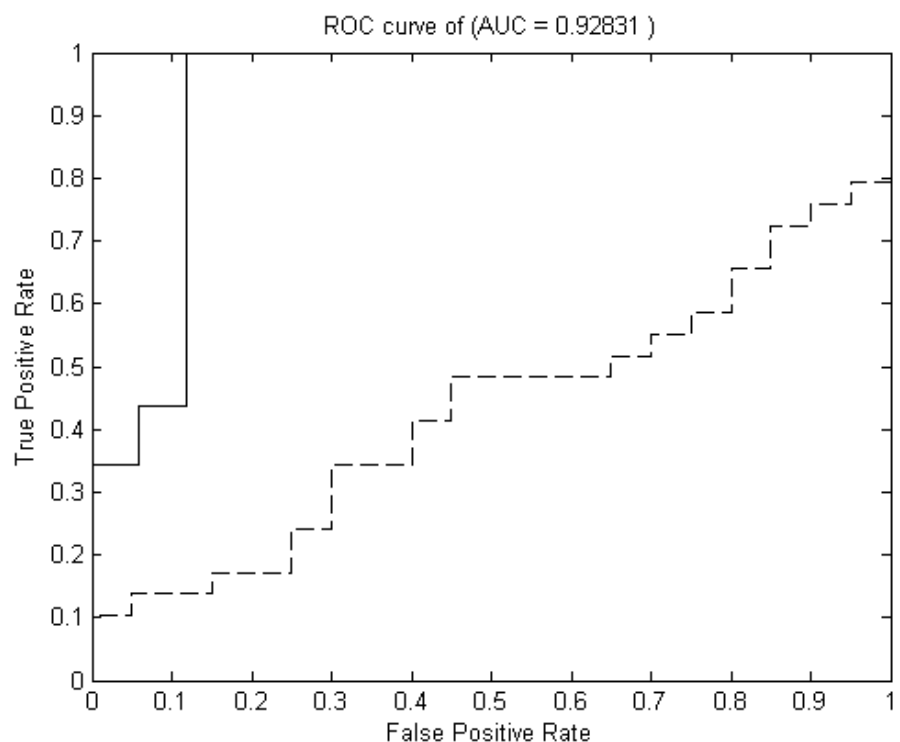
### 3.3 Area Under Curve (ROC)

En las siguientes gráficas se muestra una comparación de las curvas ROC de las que se extrajeron los valores de AUC mostrados en las tablas anteriores.



*Fig. 11. Curva ROC de la EM*

La línea discontinua representa la curva ROC de la clasificación de pacientes de EM usando el average thickness, su AUC es 0.539. La línea continua representa la clasificación usando las 9 capas con 512 puntos y el AUC es el mostrado en la parte superior de la imagen, 0.92147.



*Fig. 12. Curva ROC de la EP*

La línea discontinua representa la curva ROC de la clasificación de pacientes de EP usando el average thickness y obtiene un AUC = 0.429. La línea continua representa la clasificación usando las 9 capas con 512 puntos y el AUC es el mostrado en la parte superior de la imagen, 0.92831.

superior de la imagen, 0.92831.

En ambas gráficas se observa como las curvas se acercan mucho a un diagnostico perfecto, estando sus AUC por encima del 0.92 en ambos casos y muy por encima de las curvas del average thickness.



## Capítulo 4. Conclusiones

Como conclusión principal de este TFM puede deducirse que el nuevo prototipo de segmentación del dispositivo de Tomografía de Coherencia Óptica tipo Spectralis puede ofrecer un nuevo método diagnóstico de las enfermedades de Esclerosis Múltiple y Parkinson con una alta precisión diagnóstica, seguro, rápido e inocuo.

El proceso de validación de los datos fue clave para alcanzar los resultados finales, sin este proceso los resultados podrían haberse quedado estancados en los obtenidos con los datos originales. En esa primera prueba se obtuvieron precisiones del 23% en EM y del 60% en Parkinson. Pero después del proceso de validación, ambas precisiones superaron el 90%.

Siendo justos, los 13 pacientes que se eliminaron por contener demasiados errores, harían disminuir levemente los resultados alcanzados. Pero hay que mencionar que la máquina que extrae los datos todavía es un prototipo, la versión final podría disminuir el número de pacientes con datos erróneos. Si se llegaran a eliminar los errores de la capa 9, podrían mejorar las precisiones de las clasificaciones, sobre todo la de Parkinson, ya que esta capa parecía ser de las más importantes a la hora de diagnosticar esta enfermedad.

Por otro lado, se puede afirmar que los SVM son una herramienta muy útil para el diagnóstico de enfermedades y podrían ayudar a los médicos a manejar la gran cantidad de información que las nuevas tecnologías son capaces de extraer de nuestro organismo.

Para completar el trabajo, se realizaron varias pruebas extra. Se intentó realizar la clasificación de todos los pacientes utilizando un único SVM multiclase, con cuatro clases. La precisión en este caso bajó al 60%. También se probó con tres clases, uniendo los dos conjuntos de pacientes sanos, pero no mejoró la clasificación. Siguiendo esta tendencia, se agrupó también a los enfermos, de modo que se volvió a un clasificador binario con enfermos por un lado y sanos por otro, pero sin diferenciar su enfermedad. En este caso se volvieron a obtener precisiones en torno al 90%.

Aparentemente el SVM es capaz de distinguir entre sanos y enfermos, pero no acierta en el tipo de enfermedad. Para comprobar esto, se intentó clasificar un grupo de pacientes de Parkinson con el modelo de EM y viceversa. Los resultados volvieron a estar en torno al 90%.

Para finalizar, dados los buenos resultados obtenidos, está en proceso la publicación del contenido de este TFM en revistas de medicina.

## Capítulo 5. Trabajo futuro

Las posibilidades de trabajo futuro de este TFM son muy amplias. Las conclusiones del TFM deberán ser corroboradas con poblaciones diferentes de pacientes y, probablemente, pasar algún test doble ciego antes de poder ser aceptado como un nuevo método diagnóstico de las enfermedades tratadas.

Una vez aceptado, una opción interesante es la integración de la clasificación con SVM en el prototipo de Tomografía de Coherencia Óptica, con el objetivo de que la máquina muestre en una pantalla la probabilidad de que el paciente explorado padezca alguna de estas enfermedades. De modo que la exploración y un primer diagnóstico orientativo se realizaran en unos pocos minutos, todo usando la misma máquina y de manera totalmente automática. Incluso se habla de la posibilidad de crear una patente basada en este diagnóstico automático.

Comprobar si es efectivo en el diagnóstico precoz de estas enfermedades, es otro de los objetivos claves del equipo de investigación de médicos y también si puede ser efectivo para seguir la evolución de las enfermedades. Serán necesarias más investigaciones para poder comprobar estos puntos.

Si se comprueba la efectividad para el diagnóstico precoz y para el seguimiento de la evolución de las enfermedades, podrían surgir nuevas opciones para comprobar la efectividad de fármacos.

Expandir este mismo estudio a otras enfermedades neurodegenerativas como el Alzheimer o incluso aplicar el diagnóstico basado en redes neuronales a cualquier otro sector de la medicina, expande enormemente el abanico de posibilidades futuras de este TFM.

## **Capítulo 6. Descripción de los anexos**

En este capítulo se describe el contenido de los anexos que acompañan a la memoria de este TFM.

Anexo A. Diagrama de Gant. Muestra la evolución temporal del TFM mediante un diagrama de Gant.

Anexo B. Cálculo de los datos estadísticos. Descripción de algunos de los cálculos realizados.

Anexo C. Resultados de una clasificación. Se muestran los resultados obtenidos por los pacientes del grupo de validación en una clasificación concreta.

Anexo D. Prácticas realizadas. Descripción de las prácticas realizadas para cumplir con los objetivos del máster.

# Anexo A. Diagrama de Gant

Muestra la evolución temporal del TFM mediante un diagrama de Gant, indicando el tiempo dedicado a cada una de las tareas desarrolladas.

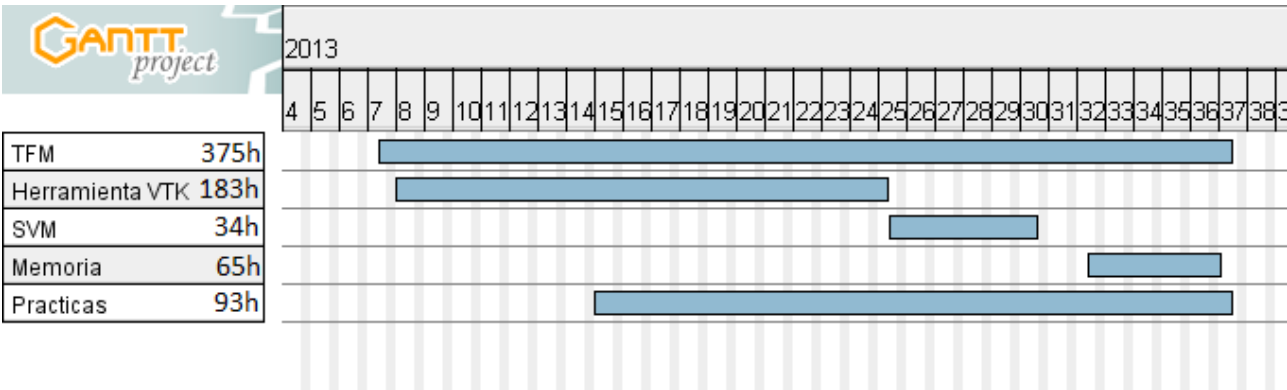


Fig 13. Diagrama de Gant

En total se han dedicado 375 horas a la realización de este TFM. La parte a la que se ha dedicado más tiempo es el desarrollo de la herramienta de visualización, incluyendo en esta tarea la validación de los datos. Pero como se ha mencionado anteriormente, no se hubieran alcanzado tan buenos resultados sin esta herramienta.

## Anexo B. Calculo de los datos estadísticos

### - 95% CI

El calculo del intervalo de confianza al 95% del AUC se realizó siguiendo el método asymptotico y se calcula como:

$$W \pm Z_{\alpha} * SE(W)$$

donde  $W$  es el AUC,  $Z_{\alpha}$  es el valor  $Z$  tal que  $(1-\alpha)$  del área bajo la distribución standard normal cae entre  $-Z$  y  $Z$  (1.96 para  $\alpha = .05$ , como ejemplo del IC al 95%), y  $SE(W)$  es el error standard del AUC.

El valor  $SE(W)$  se calculo según la referencia bibliográfica [9].

### - Sensibilidad

La sensibilidad es la probabilidad de que para un sujeto enfermo se obtenga en una prueba diagnóstica un resultado positivo.

$$Sensibilidad = \frac{VP}{VP + FN}$$

Donde  $VP$  es verdaderos positivos y  $FN$  falsos negativos.

### - Especificidad

La **especificidad** de una prueba es la probabilidad de que un sujeto sano tenga un resultado negativo en la prueba.

$$Especificidad = \frac{VN}{VN + FP}$$

Donde  $VN$ , serían los verdaderos negativos y  $FP$ , los falsos positivos.

### - Tasa de riesgo (Likelihood Ratio)

La tasa de riesgo positiva se calcula como

$$LR+ = \frac{Sensibilidad}{1 - Especificidad}$$

La Tasa de riesgo negativa se calcula como

$$LR- = \frac{1 - Sensibilidad}{Especificidad}$$

## Anexo C. Resultados de una clasificación

En este anexo se muestran los resultados obtenidos en una clasificación concreta, en la que pueden apreciarse los valores que obtiene cada paciente.

### Esclerosis Múltiple, modelo lineal

Precisión de la clasificación: 96.8750% (62/64), únicamente dos errores

Error cuadrático obtenido: 0.1250

Coefficiente de correlación al cuadrado: 0.8025

El error cuadrático es un estimador que cuantifica la diferencia entre los valores estimados y los verdaderos valores de la cantidad que se estima, es mejor cuanto menor sea.

El coeficiente de correlación al cuadrado solo se utiliza en regresiones y este no es el caso.

Probabilidades de los 64 pacientes de pertenecer al grupo de enfermos o de sanos según el modelo anterior.

*Tabla 9. Resultados de cada paciente de una clasificación con el modelo de EM*

Paciente	Enfermo	Sano	Paciente	Enfermo	Sano	Paciente	Enfermo	Sano
1	0.1816	0.8184	12	0.8817	0.1183	23	0.0053	0.9947
2	0.8851	0.1149	13	0.8914	0.1086	24	0.0000	1.0000
3	0.9064	0.0936	14	0.9129	0.0871	25	0.9178	0.0822
4	0.8164	0.1836	15	0.8969	0.1031	26	0.9273	0.0727
5	0.8995	0.1005	16	0.8963	0.1037	27	0.0138	0.9862
6	0.8672	0.1328	17	0.9326	0.0674	28	0.8695	0.1305
7	0.8712	0.1288	18	0.9147	0.0853	29	0.0329	0.9671
8	0.9154	0.0846	19	0.8537	0.1463	30	0.9367	0.0633
9	0.8916	0.1084	20	0.0395	0.9605	31	0.8534	0.1466
10	0.9282	0.0718	21	0.8646	0.1354	32	0.8694	0.1306
11	0.9396	0.0604	22	0.9222	0.0778	33	0.9604	0.0396

## Anexo C. Resultados de una clasificación

Tabla 10. Resultados de cada paciente de una clasificación con el modelo de EM (continuación)

Paciente	Enfermo	Sano	Paciente	Enfermo	Sano	Paciente	Enfermo	Sano
34	0.8582	0.1418	45	0.8917	0.1083	56	0.8898	0.1102
35	0.8206	0.1794	46	0.8449	0.1551	57	0.8984	0.1016
36	0.9390	0.0610	47	0.8987	0.1013	58	0.8694	0.1306
37	0.8840	0.1160	48	0.8205	0.1795	59	0.0397	0.9603
38	0.3600	0.6400	49	0.8968	0.1032	60	0.8999	0.1001
39	0.8979	0.1021	50	0.9223	0.0777	61	0.8416	0.1584
40	0.9715	0.0285	51	0.2104	0.7896	62	0.8450	0.1550
41	0.8808	0.1192	52	0.8596	0.1404	63	0.9203	0.0797
42	0.8884	0.1116	53	0.9166	0.0834	64	0.8904	0.1096
43	0.0463	0.9537	54	0.9289	0.0711			
44	0.8507	0.1493	55	0.8901	0.1099			

Se observa que puede diferenciarse claramente a que grupo pertenece cada uno de los pacientes. Únicamente un paciente obtiene puntuaciones inferiores al 75% (el número 38) y solo dos pacientes se han clasificado erróneamente.

Como se indica en las posibles líneas de trabajo futuro, mostrar los porcentajes obtenidos por los pacientes en la pantalla del Spectralis es una posibilidad a tener en cuenta. Los valores mostrados serían similares a los de las tablas 9 y 10. Para que estos valores puedan ayudar al diagnóstico del paciente, deberían ser lo más diferenciados posibles entre los pacientes enfermos y los sanos.

Al menos con esta población de pacientes se cumple claramente este requisito.

## Anexo C. Resultados de una clasificación

### Enfermedad de Parkinson, modelo lineal

Precisión de la clasificación: 95.9184% (47/49), únicamente dos errores

Error cuadrático obtenido: 0.1633

Coefficiente de correlación al cuadrado: 0.8304

Probabilidades de los 49 pacientes de pertenecer al grupo de enfermos o de sanos según el modelo anterior.

*Tabla 11. Resultados de cada paciente de una clasificación con el modelo de EP*

Paciente	Enfermo	Sano	Paciente	Enfermo	Sano	Paciente	Enfermo	Sano
1	0.8388	0.1612	18	0.0000	1.0000	35	0.7999	0.2001
2	0.8383	0.1617	19	0.0000	1.0000	36	0.7622	0.2378
3	0.7889	0.2111	20	0.2069	0.7931	37	0.8294	0.1706
4	0.8250	0.1750	21	0.8741	0.1259	38	0.8535	0.1465
5	0.8612	0.1388	22	0.0053	0.9947	39	0.7896	0.2104
6	0.6687	0.3313	23	0.8380	0.1620	40	0.0083	0.9917
7	0.7512	0.2488	24	0.0000	1.0000	41	0.8548	0.1452
8	0.8212	0.1788	25	0.8059	0.1941	42	0.8250	0.1750
9	0.0000	1.0000	26	0.8726	0.1274	43	0.8681	0.1319
10	0.8484	0.1516	27	0.0000	1.0000	44	0.8331	0.1669
11	0.0484	0.9516	28	0.0137	0.9863	45	0.7589	0.2411
12	0.9278	0.0722	29	0.0092	0.9908	46	0.0271	0.9729
13	0.9095	0.0905	30	0.0530	0.9470	47	0.8250	0.1750
14	0.8261	0.1739	31	0.8474	0.1526	48	0.8909	0.1091
15	0.8826	0.1174	32	0.8761	0.1239	49	0.7602	0.2398
16	0.0000	1.0000	33	0.8071	0.1929			
17	0.8638	0.1362	34	0.2139	0.7861			

En el caso de la enfermedad de Parkinson, solo un paciente obtiene resultados inferiores al 75% (el número 6).



## Anexo D. Prácticas realizadas

Para poder completar el Máster en Ingeniería Biomédica es requisito haber realizado una estancia en prácticas en algún centro Hospitalario o de I+D+I que se dedique a la Ingeniería Biomédica.

Las prácticas se han realizado en el Instituto Aragonés de Ciencias de la Salud, el personal con el que se ha colaborado durante estas prácticas es el mismo que el del TFM. Las tareas a realizar durante las prácticas han estado relacionadas con el desarrollo informático y estadístico de I+D+I de evaluación de tecnologías diagnosticas.

A continuación se describen las tareas realizadas en las prácticas.

Se ayudó a calcular el ratio estadístico PRESS<sup>[21]</sup> (predicted residual sum of squares) en el artículo con título "LDF en Parkinson vs Sanos usando protocolo axonal y maculares" cuyo objetivo principal consistía en calcular una función discriminante lineal (LDF) que permitiera distinguir entre pacientes sanos y enfermos de Parkinson.

Se realizó un estudio bastante similar al trabajo principal de este TFM en el que se comprobó la capacidad diagnostica de los SVM en la enfermedad de Glaucoma, una enfermedad de los ojos que se caracteriza por la pérdida progresiva de las fibras nerviosas del nervio óptico y cambios en su aspecto.

Las máquinas que ayudan a diagnosticar esta enfermedad muestran por pantalla, entre otros datos, el círculo ocular dividido en 6 secciones. Cada una de ellas muestra el grosor medio de la capa de fibras nerviosas del ojo en esa sección, con el objetivo de ayudar al médico a realizar el diagnóstico de la enfermedad.

Sin embargo, esta división ocular es totalmente arbitraria y los médicos no están seguros de que sea la mejor forma de realizarla, basándose en su conocimiento anatómico del ojo y en la distribución de las fibras nerviosas.

El estudio se centró en comprobar si existía una mejor forma de dividir el círculo ocular, que pudiera aportar más o mejor información para el diagnóstico del Glaucoma.

Siguiendo una metodología similar a la realizada en el TFM, se realizó la clasificación de un grupo de 240 pacientes, utilizando las mismas configuraciones de los SVM.

- Linear kernel: utiliza una función lineal.
- Precomputed kernel: utiliza una función definida por el usuario partiendo de los datos.
- RBF: utiliza una función radial similar a curvas gaussianas.
- 10-fold cross validation: validación cruzada de orden 10

En esta ocasión los datos asociados a los pacientes eran 768 puntos que indicaban el grosor

#### Anexo D. Prácticas realizadas

de la capa de fibras nerviosas a lo largo de todo el círculo ocular. A la hora de realizar las clasificaciones, se utilizaron 16 formas diferentes de agrupar los 768 puntos, para comprobar cual de estas agrupaciones obtenía mejores resultados.

En todos los casos, se dividieron los datos en 140 (aprox. 60%) para entrenar y 100 para validar, con la intención de que el SVM generalice bien. Los valores de las tablas se refieren al porcentaje de acierto en la clasificación.

Las columnas de las tablas indican el número de secciones en las que se divide el círculo ocular y si tienen alguna peculiaridad. La división en un único sector equivale a la media de los 768 puntos, 2 sectores equivale a dividir los 768 puntos en dos grupos y calcular la media de cada uno de ellos, y así sucesivamente.

Tabla 12. Precisiones dependiendo del número de secciones

	<b>1</b>	<b>2</b>	<b>4</b>	<b>6</b>	<b>8</b>	<b>12</b>	<b>16</b>	<b>24</b>	<b>32</b>	<b>64</b>	<b>768</b>
<b>Lineal</b>	79	79	77	77	77	73	72	71	72	67	65
<b>Precomputed</b>	79	79	77	77	77	73	72	71	72	67	65
<b>RBF</b>	75	82	81	77	79	72	74	77	75	72	73
<b>10-fold</b>	74,1	74,1	72,5	73,3	72,9	70,8	68,7	70,0	67,5	61,6	63,3

Tabla 13. Precisiones dependiendo del número de secciones, con secciones irregulares

	<b>2 horizontal</b>	<b>4 INST</b>	<b>4 planimetric</b>	<b>6 irregulares</b>	<b>6 irregulares horizontal</b>
<b>Lineal</b>	78	79	76	79	78
<b>Precomputed</b>	78	79	76	79	78
<b>RBF</b>	78	79	80	73	82
<b>10-fold</b>	73,75	73,33	73,33	71,66	72,5

Según estos resultados, las divisiones en 2 sectores verticales y 6 horizontales son las que mejores resultados ofrecen, llegando a un 82% de acierto. Pero, teniendo en cuenta los valores de la validación cruzada, tanto la división en un sector (valor medio), como la división en dos sectores verticales, ofrecen los mejores resultados.

La validación cruzada suele ser un método mucho más aceptado a la hora de validar resultados, pues se utilizan todos los datos disponibles para entrenar y validar.

Sería conveniente comprobar estos resultados con una base de datos de más pacientes, ya que los resultados obtenidos varían muy poco y los valores de la validación cruzada indica que probablemente dependen bastante de como se realiza la separación en pacientes para entrenar y pacientes para validar.

Para comprobar esta dependencia, se cambio el número de pacientes que se utiliza para entrenar a 163 (el 80%) y se obtuvieron unos resultados completamente diferentes.

#### Anexo D. Prácticas realizadas

Tabla 14. Precisiones dependiendo del número de secciones

	1	2	4	6	8	12	16	24	32	64	768
<b>Lineal</b>	76,62	76,62	76,62	76,62	75,32	75,32	75,32	74,02	75,32	70,12	63,63
<b>Precomputed</b>	76,62	76,62	76,62	76,62	75,32	75,32	75,32	74,02	75,32	70,12	63,63
<b>RBF</b>	72,72	76,62	79,91	75,32	76,62	74,02	74,02	72,72	75,32	75,32	71,42
<b>10-fold</b>	74,16	74,16	72,91	73,75	74,16	72,91	72,91	70,83	71,66	63,75	63,75

Tabla 15. Precisiones dependiendo del número de secciones, con secciones irregulares

	2 horizontal	4 INST	4 planimetric	6 irregulares	6 irregulares horizontal
<b>Lineal</b>	76,62	75,32	76,62	75,32	75,32
<b>Precomputed</b>	76,62	75,32	76,62	75,32	75,32
<b>RBF</b>	72,72	80,51	76,62	77,92	79,22
<b>10-fold</b>	73,75	73,75	74,16	71,66	72,91

En base a estos resultados, resulta difícil decidir que separación es la mejor y el resultado de este estudio termino por no ser concluyente.

Partiendo de los resultados obtenidos al usar 768 puntos, se intento seleccionar los puntos con mayor relevancia<sup>[17]</sup>, con el objetivo de observar cuales son los sectores más relevantes a la hora de clasificar con un SVM e intentar diseñar una nueva forma de dividir el círculo ocular en base a estos resultados.

Con los puntos más relevantes extraídos, se volvió a realizar la clasificación y se dibujaron para comprobar a que sectores del ojo pertenecían.

Tabla 16. Precisiones usando los puntos más significativos

	Pesos >0.25 = 156 puntos	> 0.3 = 92	> 0.4 = 33	> 0.45 = 14
<b>Lineal</b>	64	66	67	73
<b>Precomputed</b>	64	66	67	73
<b>RBF</b>	61	60	70	75
<b>10-fold</b>	61,66	61,66	64,17	67,5

#### Anexo D. Prácticas realizadas

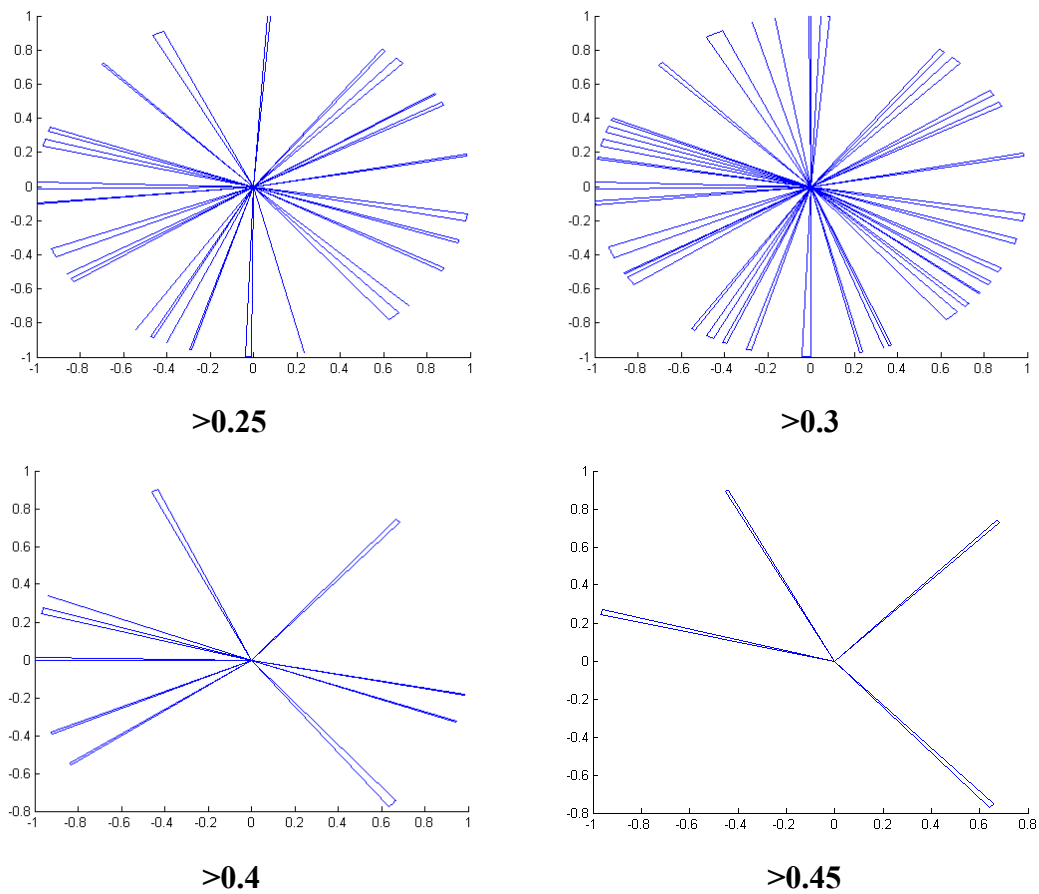


Fig 14. Posición de los puntos más significativos a lo largo del círculo ocular

A raíz de estos resultados, se observa que todas las clasificaciones son mejores que utilizando los 768 puntos y bastan 14 puntos para obtener una precisión que ronda el 75%. Aun así, no se consiguen mejores resultados que con las otras divisiones

El mismo procedimiento se repitió en los datos extraídos por otra máquina, en esta ocasión la base de datos es de 140 pacientes y 256 puntos por paciente. Los pacientes son diferentes a los del estudio anterior.

En todos los casos, se dividió los datos en 100 (aprox. 70%) para entrenar y 40 para validar. Los valores se refieren al porcentaje de acierto en la clasificación.

Tabla 17. Precisiones dependiendo del número de secciones

	1	2	4	6	8	12	16	24	32	64	256
<b>Lineal</b>	75	81,81	79,5	86,3	86,3	81,81	75	77,2	84	77,2	81,81
<b>Precomputed</b>	75	81,81	79,5	86,3	86,3	81,81	75	77,2	84	77,2	81,81
<b>RBF</b>	77,2	72	81,81	81,8	79,5	86,3	65,9	72,7	81,8	77,2	84
<b>10-fold</b>	81,9	79,8	77	82,6	79,1	81,9	79,1	79,1	79,8	84	83,3

Tabla 18. Precisiones dependiendo del número de secciones, con secciones irregulares

	2 horizontal	4 INST	4 planimetric	6 irregulares	6 irregulares horizontal
<b>Lineal</b>	86,3	79,5	86,3	81,3	88,6
<b>Precomputed</b>	86,3	79,5	86,3	81,8	88,6
<b>RBF</b>	81,8	81,8	81,8	86,3	88,6
<b>10-fold</b>	81,25	83,3	82,6	82,6	81,9

Los mejores resultados se obtienen en los 6 sectores horizontales con un 88%, pero hay muchas otras configuraciones que se aproximan mucho a ese valor, 6 sectores regulares, 8 sectores, 12 sectores, 2 horizontales, 4 planimetric y 6 irregulares.

El valor de la validación cruzada también es muy similar en todos ellos.

Probablemente sería conveniente realizar estas comprobaciones con una base de datos más grande, ya que los resultados obtenidos varían muy poco y es difícil decidir cual es la mejor configuración.

Se probó a extraer los puntos más representativos, pero en este caso la precisión empeoraba con respecto a los 256 puntos.

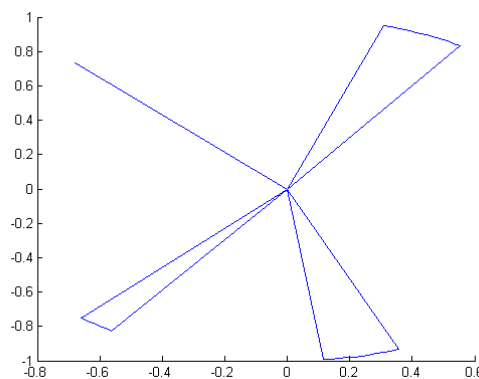


Fig 15. Posición de los puntos más significativos a lo largo del círculo ocular

## Bibliografía

- [1] Chih-Chung Chang and Chih-Jen Lin, LIBSVM : a library for support vector machines. ACM Transactions on Intelligent Systems and Technology, 2:27:1--27:27, 2011. Software available at <http://www.csie.ntu.edu.tw/~cjlin/libsvm>
- [2] Elena Garcia-Martin PhD, Luis E Pablo Prof, Jose M. Larrosa et al. Neural networks to identify multiple sclerosis with optical coherence tomography
- [3] Elena Garcia-Martin, Luis E Pablo, Raquel Herrero et al. Diagnostic ability of a linear discriminant function for Spectral domain optical coherence tomography in multiple sclerosis patients
- [4] Elena Garcia-Martin PhD, Maria Satue MD, Sofia Otin MD et al. Retina measurements for diagnosis of Parkinson disease
- [5] Elena Garcia-Martin PhD, Luis E. Pablo PhD, Raquel Herrero et al. Diagnostic Ability of a linear discriminant function for spectral-domain optical coherence tomography in patients with multiple sclerosis
- [6] Larrosa JM, Polo V, Ferreras A et al. Neural network analysis of different segmentation strategies of nerve fiber layer assessment for glaucoma diagnosis
- [7] Chih-Wei Hsu, Chih-Chung Chang, and Chih-Jen Lin, A Practical Guide to Support Vector Classification. 8 Sept. 2013  
<<http://www.csie.ntu.edu.tw/~cjlin/papers/guide/guide.pdf>>
- [8] Douglas G Altman, ROC curves and confidence intervals: getting them right. 8 Sept. 2013  
<<http://www.ncbi.nlm.nih.gov/pmc/articles/PMC1729299/pdf/v083p00236.pdf>>
- [9] Lara Hopley, Jo van Schalkwyk, The magnificent ROC. 8 Sept. 2013  
<<http://www.anaesthetist.com/mnm/stats/roc/Findex.htm>>
- [10] López de Ullibarri Galparsoro I, Pita Fernández, S., Curvas ROC Unidad de Epidemiología Clínica y Bioestadística. Complejo Hospitalario Juan Canalejo. A Coruña (España), CAD ATEN PRIMARIA 1998; 5 (4): 229-235., 8 Sept. 2013 <[http://www.fisterra.com/mbe/investiga/curvas\\_roc/curvas\\_roc.htm](http://www.fisterra.com/mbe/investiga/curvas_roc/curvas_roc.htm)>
- [11] Wen Zhu, Nancy Zeng, Ning Wang, Sensitivity, Specificity, Accuracy, Associated

## Bibliografía

- Confidence Interval and ROC Analysis with Practical SAS Implementations,  
[http://www.floppybunny.org/robin/web/virtualclassroom/dss/articles/sensitivity\\_specificity\\_accuracy\\_CI.pdf](http://www.floppybunny.org/robin/web/virtualclassroom/dss/articles/sensitivity_specificity_accuracy_CI.pdf)
- [12] Web Gabriel Rada. Revisado 2007 Tomás Merino, Intervalos de confianza , 8 Sept. 2013 <<http://escuela.med.puc.cl/Recursos/recepidem/EPIANAL9.HTM>>
- [13] The Visualization Toolkit (VTK), 8 Sept. 2013, <<http://www.vtk.org/>>
- [14] K. Kampa, libsvm for MATLAB tutorial, 8 Sept. 2013, <[https://sites.google.com/site/kittipat/libsvm\\_matlab](https://sites.google.com/site/kittipat/libsvm_matlab)>
- [15] FELEM, organización sin ánimo de lucro declarada de Utilidad Pública, 8 Sept. 2013, <<http://www.esclerosismultiple.com/>>
- [16] Viatusalud, Clínica Universitaria de Navarra, España., 8 Sept. 2013, <[http://geosalud.com/adultos\\_mayores/parkinson.htm](http://geosalud.com/adultos_mayores/parkinson.htm)>
- [17] Yin-Wen Chang, Chih-Jen Lin, Feature Ranking Using Linear SVM, 8 Sept. 2013, <<http://www.csie.ntu.edu.tw/~cjlin/papers/causality.pdf>>
- [18] Sergio Herrera Alonso, Breve manual Tcl, 8 Sept. 2013, <[http://www-gris.det.uvigo.es/~estela/SC0708/breve\\_manual\\_tcl.pdf](http://www-gris.det.uvigo.es/~estela/SC0708/breve_manual_tcl.pdf)>
- [19] Tcl Reference Manual: Contents, 8 Sept. 2013, <<http://tmml.sourceforge.net/doc/tcl/index.html>>
- [20] Oracle Data Mining Concepts, 8 Sept. 2013, <[http://docs.oracle.com/cd/B28359\\_01/datamine.111/b28129/algo\\_svm.htm#CHDBDHFB](http://docs.oracle.com/cd/B28359_01/datamine.111/b28129/algo_svm.htm#CHDBDHFB)>
- [21] Does SPSS produce the PRESS statistic?, 8 Sept. 2013, <<http://www-01.ibm.com/support/docview.wss?uid=swg21479373>>
- [22] Cross-validation (statistics), 8 Sept. 2013, <[http://en.wikipedia.org/wiki/Cross-validation\\_%28statistics%29](http://en.wikipedia.org/wiki/Cross-validation_%28statistics%29)>
- [23] Validación cruzada, 8 Sept. 2013, <[https://es.wikipedia.org/wiki/Validaci%C3%B3n\\_cruzada](https://es.wikipedia.org/wiki/Validaci%C3%B3n_cruzada)>
- [24] Likelihood ratios in diagnostic testing, 8 Sept. 2013, <[http://en.wikipedia.org/wiki/Likelihood\\_ratios\\_in\\_diagnostic\\_testing](http://en.wikipedia.org/wiki/Likelihood_ratios_in_diagnostic_testing)>
- [25] Curva ROC, 8 Sept. 2013, <[http://es.wikipedia.org/wiki/Curva\\_ROC](http://es.wikipedia.org/wiki/Curva_ROC)>