

UNIVERSIDAD AUTÓNOMA DE MADRID

ESCUELA POLITÉCNICA SUPERIOR



TRABAJO FIN DE MÁSTER

Implementación de una herramienta de aprendizaje automático como apoyo para el diagnóstico temprano de enfermedades neurodegenerativas

Máster Universitario en Bioinformática y Biología Computacional

Autora: Couce Iglesias, Marta

Tutor: Varona Martínez, Pablo

Departamento de ingeniería informática

Junio de 2021

Índice general

Agradecimientos	4
Abstract.....	5
Resumen	6
1. Estado del arte	7
1.1 Tecnología de seguimiento ocular	7
1.1.1 Técnicas de seguimiento ocular	8
1.1.2 Estimación de la mirada en el seguimiento ocular	9
1.1.3 Análisis del movimiento ocular	12
1.2 Enfermedades neurodegenerativas	13
1.2.1 Enfermedad de Alzheimer (EA)	13
1.2.2 Enfermedad de Parkinson (EP).....	14
1.2.3 Demencia frontotemporal (DF).....	15
1.2.4 Demencia con cuerpos de Lewy (DCL)	16
1.3 Neuroanatomía de la visión en enfermedades neurodegenerativas.....	16
1.3.1 Movimientos oculares en Alzheimer.....	19
1.3.2 Movimientos oculares en la Enfermedad de Parkinson.....	20
1.3.3 Movimientos oculares en la Demencia Frontotemporal.....	20
1.3.4 Movimientos oculares en Demencia por Cuerpos de Lewy	21
2. Diseño y desarrollo	22
2.1 Fase 1: Elección de los dispositivos.....	22
2.2 Fase 2: Diseño del estudio	26
2.3 Fase 3: Obtención de los datos	29
2.4 Fase 4: Elección de los métodos de análisis de datos	31
2.4.1 Long Short-Term Memory (LSTM)	32

2.4.2 Multivariate Time Series Classification with Learned Discretization (SMTS)	34
3. Resultados y análisis.....	37
3.1 Resultados de la prueba de eye-tracking.....	37
3.2 Comparativa de los algoritmos de clasificación	40
3.2.1 SMTS	40
3.2.1.1 Optimización de SMTS	42
3.2.2 LSTM	43
3.3 Clasificación por tipos de enfermedad.....	45
4. Conclusiones	47
5. Trabajo futuro	49
Bibliografía	51
ANEXO I.....	55
ANEXO II.....	56
ANEXO III.....	61

Agradecimientos

Me gustaría agradecer todo el tiempo y el trabajo que ha llevado este pequeño proyecto de investigación a la gente que me ha apoyado y ha estado ahí a lo largo de estos meses.

A mis compañeros de piso, Carlos, Ana y Pellumb. Gracias por las risas, los consejos y las largas charlas que me ayudaron a pasar los momentos más difíciles cuando parecía que no iba a salir adelante.

Gracias a Dani, por siempre estar ahí para mí, donde sea y cuando sea.

Gracias a mis padres y a mi hermana, que cuando ven que las situaciones me superan me dan fuerzas para mirar más allá y ver la perspectiva en la vida.

Gracias a todas mis amigas con las que he podido compartir las alegrías y triunfos y las penas y frustraciones que me han acompañado realizando y escribiendo este trabajo.

Gracias a Pablo, mi tutor.

Abstract

Eye tracking technology is used in all sorts of areas to establish a relationship between cognitive function and eye movement: from marketing in large companies, to behavioural studies in psychology and a large variety of biomedical applications.

The diversity in modes of operation of the devices, and the treatment of time series that each one offers as output, result in a complex setback in the use of this technique. Therefore, the objective of this work is the creation and validation of a simple protocol of support in the diagnosis of neurodegenerative diseases from the complete time series resulting from eye monitoring to avoid loss of information.

The data were obtained from the collaboration with a nursing home and an hospital. To generate the time series from the reading of small texts a previously developed application of intuitive and simple operation was used. For this phase, due to its technical specifications, the Tobii X2-30 eye-tracker was chosen for data collection.

In this project, two machine learning algorithms were selected to make a comparison between them and, finally, use the most accurate one to support medical diagnosis. Multivariate Time Series Classification with Learned Discretization (SMTS) and a Long Short Term Memory (LSTM) neural network were used.

After the pre-processing of the data, and the comparison of the machine learning models, a precision of 92.30% was obtained when differentiating patients with neurodegenerative disease from those who are healthy and an accuracy of 89.6% to assign the specific type of disease to each patient.

Therefore, although the results are influenced by the use of a small cohort, this study gives us a promising perspective regarding the use of this technology to support the diagnosis of neurodegenerative diseases, using a quick and simple protocol.

Resumen

La tecnología de seguimiento ocular es utilizada en todo tipo de ámbitos para establecer una relación entre la función cognitiva y el movimiento ocular: desde el marketing en grandes empresas, hasta los estudios conductuales en psicología y una enorme variedad de aplicaciones biomédicas.

La diversidad en modos de funcionamiento de los aparatos, y el tratamiento de las series temporales que ofrece cada uno como resultado, resultan en un complejo escoyo en el uso de esta técnica. Por lo tanto, el objetivo de este trabajo es la creación y validación de un protocolo sencillo de apoyo en el diagnóstico de enfermedades neurodegenerativas a partir de las series temporales completas resultantes del seguimiento ocular para evitar la pérdida de información.

Los datos se obtuvieron de la colaboración con una residencia y un hospital. Para generar las series temporales a partir de la lectura de pequeños textos se utilizó una aplicación previamente desarrollada de funcionamiento intuitivo y sencillo. Para esta fase, debido a sus especificaciones técnicas, se escogió el Eye-tracker Tobii X2-30 para la recogida de datos.

En este proyecto se usaron dos algoritmos de aprendizaje automático para hacer una comparativa entre ellos y usar el de mayor precisión para el apoyo en el diagnóstico médico. Se utilizaron los algoritmos Multivariate Time Series Classification with Learned Discretization (SMTS) y una red neuronal Long Short Term Memory (LSTM).

Tras el preprocesamiento de los datos, y la comparativa de los modelos de *machine learning*, se obtuvo una precisión de un 92,30% al diferenciar pacientes con enfermedad neurodegenerativa de aquellos sanos y una precisión de un 89,6% para asignar el tipo de enfermedad específica a cada paciente.

Por consiguiente, aunque los resultados finales estén influenciados por el uso de una cohorte pequeña, este estudio nos da una prometedora perspectiva respecto al uso de esta tecnología como apoyo en el diagnóstico de enfermedades neurodegenerativas, mediante un protocolo rápido y sencillo.

1

Estado del arte

1.1 Tecnología de seguimiento ocular

EL origen de esta tecnología data del siglo XVIII cuando Charles Bell relacionó el control del movimiento ocular con la actividad cerebral, clasificó los movimientos oculares y describió el efecto del movimiento ocular sobre la orientación visual [1].

Dichos estudios sentaron un precedente claro en el desarrollo de técnicas durante los siguientes siglos guiadas por la necesidad de aumentar el conocimiento que se tenía de las funciones cerebrales. Así pues, distintos métodos fueron surgiendo, entre ellos, Delabarre (1898) desarrolló un sistema de registro de movimientos oculares mediante plumas unidas mecánicamente a un anillo de yeso colocado sobre la córnea; Buswell (1935) capturaba la reflexión procedente de la iluminación de la córnea con un haz de luz sobre una película a través del uso de prismas creando, así, el primer procedimiento no intrusivo; Yarbus (1967) fabricó una lente corneal, que se utilizó en sus experimentos clásicos sobre la percepción y Remmel (1984) y Kenyon (1985) crearon una bobina magnética que insertada en el ojo y, conectada a un dispositivo mecánico o óptico, medía la variación de un campo electromagnético que se forma durante el movimiento ocular [1] [2].

En un inicio, los métodos utilizados eran intrusivos, caros y requerían de un especialista para obtener resultados reproducibles. Gracias al desarrollo de la tecnología, actualmente existen dispositivos *user-friendly*, baratos y no intrusivos que proporcionan resultados altamente precisos [3]. Como resultado, en los últimos 20 años se ha producido un aumento exponencial del uso de la tecnología de seguimiento ocular en todo tipo de disciplinas [4] [5] [6] [7] [8] [9] (ver Figura 1).

Uso de la tecnología de eye-tracking en investigación a lo largo de los años

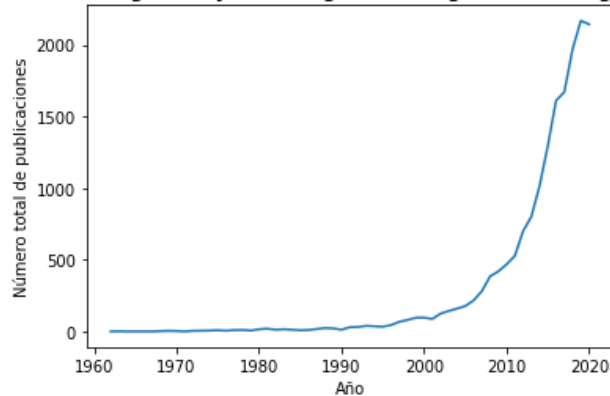


Figura 1. Número total de publicaciones entre los años 1962 y 2020 utilizando los términos de búsqueda: (“eye tracking” OR “eye-tracking” OR “eyetracking”) en Web of Science.

El ámbito de la investigación médica ha estado presente siempre en la creación y desarrollo de esta herramienta. Inicialmente, se centró en la relación más clara y directa entre el movimiento ocular y la capacidad lectora del individuo. Con el desarrollo del movimiento conductista en psicología experimental, se desarrolló el boom del estudio entre comportamientos prácticos en conducta y el movimiento ocular, estudios en los que se basaron para el uso de estos sistemas en otras áreas como el marketing. Finalmente, el desarrollo de la psicología cognitiva a partir de 1970 propulsó un análisis más centrado en la relación entre la percepción visual y los procesos mentales, sentando las bases para fomentar el uso de esta técnica en el estudio de enfermedades cognitivas [10].

1.1.1 Técnicas de seguimiento ocular

A lo largo de las décadas, se han ido desarrollando equipos cada vez más sofisticados para la medición del movimiento ocular [2] [11] [12] [13]:

1. Lentillas esclerales: consisten en pequeñas bobinas de alambre integradas en lentillas de gran tamaño que generan un potencial al verse afectadas por el campo eléctrico que se crea en el propio ojo al moverse. Son altamente precisas, pero también muy intrusivas (requieren de anestesia para su colocación) y la medición del movimiento ocular está relacionado con la posición de la cabeza lo que las hace inadecuadas para la medición del punto de visión.
2. Electrooculografía: en esta técnica se posicionan electrodos en la piel alrededor de los ojos para medir las pequeñas diferencias de potencial que se producen durante el movimiento ocular y así deducir la posición del ojo. Es un método fácil, barato y que puede calcular la posición ocular con los ojos cerrados. En contrapartida, es un

método intrusivo y, por lo tanto, ineficiente para estudios de larga duración por la incomodidad producida al individuo y, además, es dependiente de la posición de la cabeza no siendo adecuado para la medición del punto de visión.

3. Oculografía infrarroja: se mide la luz infrarroja que es reflejada por la esclera calculándose así la posición del ojo. La fuente de luz y los sensores se suelen colocar en unas gafas haciendo al método intrusivo. Es muy sensible a los cambios de luz externa y sólo puede medir movimientos 35 grados a lo largo del eje horizontal y 20 grados a lo largo del eje vertical. Como ventaja puede medir movimientos oculares en la oscuridad.
4. Videoculografía: en esta tecnología se usa el destello que es grabado por una cámara al iluminar la pupila con luz para localizar la posición del ojo. Es la técnica actual más utilizada ya que es la única que, además, proporciona información sobre el punto observado. Este método puede ser invasivo o no invasivo y, cada uno de ellos puede utilizar luz visible o infrarroja cercana de aproximadamente 880nm de longitud de onda, siendo la segunda la más comúnmente utilizada. Dentro de este sistema podemos distinguir entre:
 - *Eye Tracker* de una cámara: el principal inconveniente es el limitado campo de visión disponible para capturar imágenes con alta resolución. Este problema mejora añadiendo más fuentes de luz y se han conseguido altas precisiones con un sistema basado en una cámara centrada en la pantalla y cuatro haces de luz infrarroja cercana en las esquinas.
 - *Eye Tracker* multicámara: con esta tecnología se emplean multitud de cámaras para capturar tanto un campo de visión amplio, y así permitir el movimiento libre de la cabeza, como un campo de visión limitado para capturar imágenes de alta resolución y conseguir estimaciones fiables de la mirada. Se utilizan diversos sistemas de cámaras (cámaras individuales para cada ojo, cámaras para realizar seguimiento del movimiento de la cabeza, etc.) que combinan la información obtenida para realizar el cálculo final del punto de mira.

1.1.2 Estimación de la mirada en el seguimiento ocular

En este apartado nos centraremos en el cálculo de la trayectoria de la mirada en la tecnología más utilizada, la videoculografía.

Una vez obtenida la imagen del reflejo corneal se pueden utilizar dos métodos para el cálculo de la posición de la mirada [14] [15] [16]:

- Estimación de la mirada basada en características: identifica características locales informativas del ojo como los píxeles negros correspondientes a la pupila o al iris y reflejos de la córnea al ser iluminada con luz infrarroja (imágenes de Purkinje, se suele utilizar la primera ya que es la más brillante y fácilmente medible), ya que son generalmente menos sensibles a las variaciones en la iluminación y en el punto de fijación de la mirada. Esta tecnología tiene problemas de rendimiento bajo una fuerte luz ambiental y de estimación de la mirada cuando no se disponen de características precisas del iris y la pupila. Existen dos tipos de estimaciones:
 - Las basadas en modelos geométricos (Figura 2): utilizan un modelo geométrico explícito del ojo para estimar el vector de dirección de la mirada en 3D. La mayoría se basan en información métrica y requieren, por lo tanto, la calibración de la cámara y un modelo geométrico global de fuentes de luz y posición y orientación de la cámara y el monitor. La estrategia que siguen estos modelos es la siguiente:
 1. Reconstrucción del eje óptico: se estima a partir del centro de la córnea y de la pupila.
 2. Reconstrucción del eje visual: se calcula como un vector desde el centro del globo ocular hasta el centro del iris.
 3. Estimación del punto de la mirada: se realiza a través de la intersección del eje visual (vector de dirección de la mirada) con la geometría de la escena (objetos).

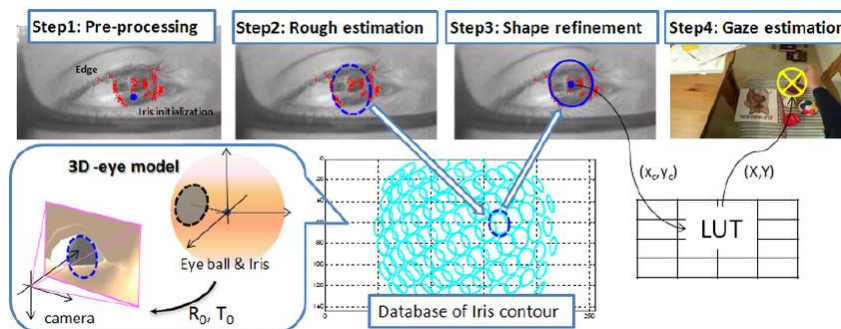


Figura 2. Método de estimación de la mirada [15]. © [2012] IEEE

- Estimaciones basadas en interpolación: no modelan explícitamente el ojo humano si no que asocian el punto observado a una función genérica (paramétrica o no paramétrica) basada en las características de la imagen. La calibración se utiliza para calcular los coeficientes desconocidos mediante ajuste numérico en las funciones paramétricas y en las no paramétricas se utilizan típicamente redes neuronales que realizan un mapeo a partir de los datos de coordenadas de ciertos puntos faciales para obtener las coordenadas de la mirada.

- Estimación de la mirada basada en la apariencia: utilizan las coordenadas de pantalla de las imágenes de los ojos para estimar la dirección de la mirada mediante una función. Las más utilizadas son:
 - Redes neuronales multicapa.
 - Interpolación lineal.
 - Distintos tipos de regresiones: lineal, gaussiana o *Random Forest*.
 - *Neighborhood Approximation Forests*.

No requiere calibración ya que el mapeo se hace directamente sobre las imágenes.

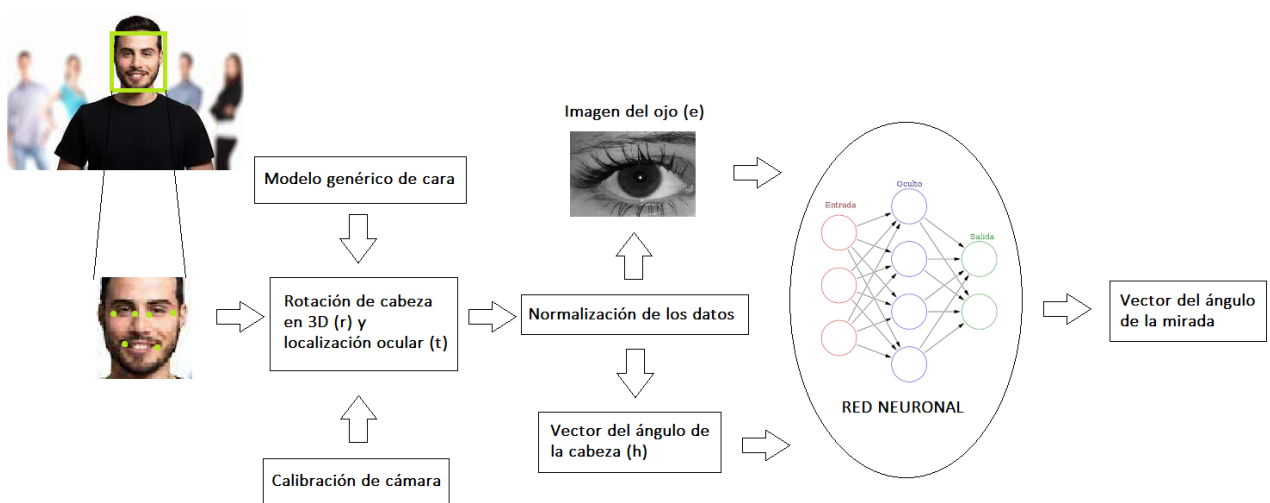


Figura 3. Método para la estimación de la mirada en la naturaleza basada en la apariencia usando redes neuronales convolucionales multimodales [16].

La generalización de este método se ve influenciada por los siguientes factores [13]:

- La variación en la apariencia física. Las diferentes combinaciones de iluminación, color de piel, textura y forma ocular, gafas, etc. pueden variar el resultado en cada individuo.
- Variación de la posición de la cabeza. Su influencia es inferior ya que se puede modelar al ser resultado de una rotación geométrica y un movimiento conocidos y existir buenas formulaciones analíticas.
- Ajuste excesivo de la estimación de puntos al tratarse de entornos con grandes variaciones para los cuales el cálculo de un set óptimo de parámetros por parte de los algoritmos es insuficiente.

1.1.3 Análisis del movimiento ocular

Generalmente, en la mayor parte de los estudios, en vez de analizar todos los datos brutos obtenidos del *eye-tracker*, estos se preprocesan para obtener los elementos básicos del movimiento ocular: fijaciones y sacadas, principalmente, junto con búsqueda visual y fluida y cognición social [2] [17].

Una fijación se define como el momento en el que el ojo está fijado en un objetivo, la percepción es estable y se está obteniendo información de la escena. Cada fijación se puede definir por cuatro atributos: tiempo de inicio, duración (aproximadamente 150-330 ms), coordenadas en x e y y dispersión espacial de las grabaciones (0.5-2°). En la realización de una fijación están involucradas múltiples zonas cerebrales como la corteza fronto-insular, la corteza cingulada anterior, el área motora suplementaria, el colliculli superior, el tálamo y la corteza prefrontal dorsolateral bilateral. El tiempo y la frecuencia de fijación se miden en algunos estudios para predecir el deterioro visual y de la atención lo que está asociado a daños en las cortezas cingulada anterior y frontal. Por otro lado, la precisión en la búsqueda visual se analiza para evaluar la disfunción ejecutiva, y las emociones faciales para reconocer una tarea de cognición social al realizarla. Durante las fijaciones el ojo nunca se encuentra quieto del todo existiendo movimientos de temblor, desviaciones y microsacadas que se pueden registrar con los dispositivos de seguimiento ocular.

Las sacadas son movimientos cortos y rápidos entre fijaciones. Como atributos definitorios están: la duración (30-80 ms), amplitud (4-20°), velocidad (30-500°/s) y aceleración (4000-8000°/s²). Dentro de este tipo de movimientos oculares se pueden diferenciar prosacadas (movimiento voluntario o automático de los ojos hacia un objetivo), antisacadas (movimiento contrario a un objetivo visual periférico) y sacadas recordadas (movimiento sacádico hacia un objetivo visual periférico que ya no está). Al realizar tareas sacádicas, el tiempo de respuesta, la latencia para cambiar entre prosacadas y antisacadas y la tasa de error antisacádica suelen ser indicativos de deterioro de la función ejecutiva debido a disfunciones corticales y subcorticales.

Estos eventos son detectados mediante diversos métodos, entre los más usados:

- Algoritmos basados en la dispersión: definen una fijación como un grupo de puntos dentro del cual la distancia entre un punto y el siguiente es más pequeña que un umbral de dispersión predefinido.
- Algoritmos basados en la velocidad: se computa la velocidad entre puntos teniendo en cuenta la distancia. Si es menor al umbral se define el punto como una fijación, y si es mayor, como una sacada.

A partir de las fijaciones, analizando su número, velocidad y duración, podemos inferir donde se fija la atención del individuo y el tiempo de estimulación de dicha zona. De las sacadas, prestando atención a su ángulo y las rutas de escaneo, podemos deducir cambios en la atención y función ejecutiva impulsados por estímulos y aquellos que son automáticos [17].

1.2 Enfermedades neurodegenerativas

Las enfermedades neurodegenerativas conforman un grupo de dolencias cuyo factor común es la atrofia progresiva y/o muerte neuronal y, por lo tanto, la pérdida de funciones motoras, sensoriales y/o cognitivas.

Son enfermedades que, actualmente, no tienen cura y se enmarcan como debilitantes. Además, son reconocidas como una de las diez principales causas de muerte a nivel mundial [18].

De entre los más de 100 síndromes que forman parte de este conjunto, las enfermedades que afectan al movimiento ocular más conocidas y con mayor tasa de prevalencia, y, por lo tanto, con mayor impacto social son [19]:

- Enfermedad de Alzheimer.
- Enfermedad de Parkinson y otros parkinsonismos.
- Demencia frontotemporal.
- Demencia con cuerpos de Lewy.

Cada una de ellas afecta a una región del sistema nervioso central que, debido a las conexiones entre neuronas, se ve ligada al sistema nervioso oculomotor de forma directa o indirecta.

1.2.1 Enfermedad de Alzheimer (EA)

Es la enfermedad neurodegenerativa de mayor prevalencia (entre un 2% y un 8% en población mayor de 60 años), representa la séptima causa de muerte mundial y conforma entre el 60% y 80% del total de casos de demencias [18] [20]. Actualmente, se estima que unos 50 millones de personas sufren esta enfermedad y son diagnosticados 8 millones de casos nuevos al año, 1 persona cada 4 segundos [20].

Esta enfermedad se caracteriza por lenguaje errático, pérdida de memoria, cambios de comportamiento, dificultad para orientarse en espacio, lugar y tiempo, pérdida de autonomía y dificultades motoras. Esta sintomatología sigue un lento gradiente ascendente que puede derivar en coma y muerte para el paciente.

Fisiopatología

El diagnóstico definitivo de la EA se realiza postmortem al observar una atrofia simétrica de los lóbulos temporales, frontales, parietales u occipitales, siendo el primero el más frecuente. Esta atrofia deriva en el aumento de la profundidad de los surcos, la disminución del espesor de las circunvoluciones, del peso y del volumen cerebral y la dilatación del

sistema ventricular. Otra alteración macroscópica observable es la atrofia selectiva del hipocampo acompañada de dilatación del asta temporal del ventrículo lateral [21] [22].

Respecto a las lesiones microscópicas, destacan las placas neuríticas y los ovillos neurofibrilares. Las placas neuríticas se forman a partir de fibras de péptido amiloideo (A β) y dan lugar a un efecto inflamatorio, interfieren en la sinapsis neuronal, dañan los sistemas GABAérgicos y glutaminérgicos y actúan como neurotoxinas. Estas placas están altamente concentradas en el núcleo estriado piramidal, donde encontramos las dendritas basales y apicales de las células piramidales de CA1. Esta región del hipocampo es crítica en la orientación espacial, en el aprendizaje y en diversas funciones relacionadas con la memoria [21] [22] [23].

Los ovillos neurofibrilares son acúmulos citoplasmáticos de parejas de fibrillas de 10 nm de diámetro que se disponen de forma helicoidal formando ovillos, asas o bobinas debido a la hiperfosforilación de la proteína tau. La mayor concentración de fibras se encuentra, al igual que con las placas, en la zona CA1 del hipocampo [21] [22] [23].

1.2.2 Enfermedad de Parkinson (EP)

La enfermedad de Parkinson (EP) es la segunda enfermedad neurodegenerativa de mayor prevalencia tras la EA con un impacto socioeconómico de gran peso.

Es un trastorno con un período de latencia preclínico de entre 5 y 10 años y que se caracteriza por la presencia de síntomas motores y no motores. Dentro de los primeros, los principales serían el temblor en reposo de una frecuencia de 4-6 Hz asimétrico y distal en las extremidades superiores (se da en el 50-70% de los casos), la rigidez, la bradicinesia y la inestabilidad postural que aparece a medida que evoluciona la enfermedad llegando a una imposibilidad de iniciar la marcha o a un notable titubeo ante un obstáculo. Respecto a los síntomas no motores, muchos de ellos aparecen previa o simultáneamente a las manifestaciones motoras. Los más representativos de los últimos son trastornos sensitivos (afectan al 30-50% de los pacientes y comprenden dolor, frialdad, hormigueo y entumecimiento), disfunción del sistema nervioso autónomo (afecta al 90% de los enfermos, siendo la manifestación más importante la hipotensión ortostática), demencia (40% de afectados), depresión (20-30%), trastornos de ansiedad, alucinaciones y psicosis (estas dos últimas por los fármacos administrados para su tratamiento) [24] [25].

Fisiopatología

La EP se caracteriza por una pérdida creciente de neuronas dopaminérgicas en la sustancia nigra pars compacta del mesencéfalo. Para que se desarrolle la enfermedad dicha pérdida tiene que ser superior al 50%, de tal forma que los niveles de dopamina disminuyan más de un 80% [24].

Esto junto a su otra característica propia, los cuerpos de Lewy (agregados insolubles de la proteína α -sinucleína), da lugar a la denervación dopaminérgica de las proyecciones desde la sustancia nigra pars compacta hacia el núcleo estriado, lo que deriva en la alteración de los ganglios basales, que se relaciona con las principales manifestaciones de la EP [24] [25].

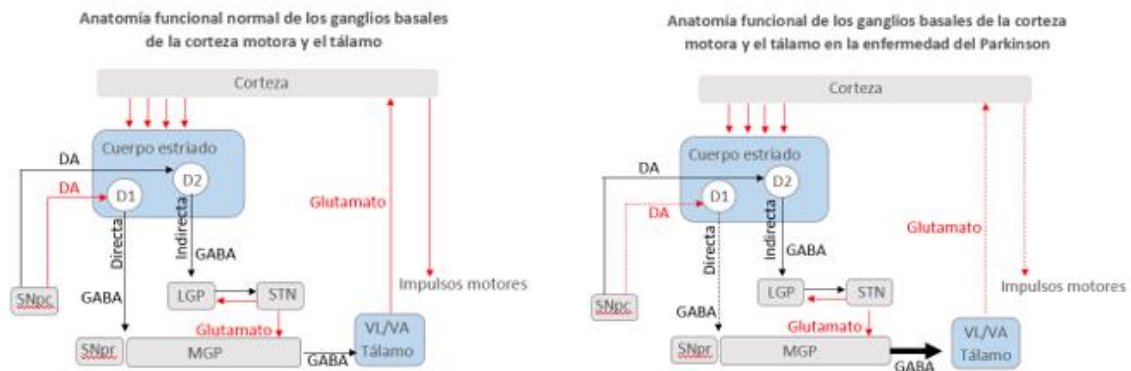


Figura 4. Esquema comparativo del circuito neuronal normal y en la enfermedad de Parkinson. DA: dopamina; SNpc: sustancia nigra pars compacta; SNpr: sustancia nigra pars reticulata; LGP: globus pallidus lateral; STN: núcleo subtalámico; MGP: globus pallidus medial. En rojo los transmisores excitadores (Glutamato, DA y D1) y en negro los inhibidores (GABA, DA, D2).

1.2.3 Demencia frontotemporal (DF)

La DF es la causa más frecuente de demencia en personas menores de 60 años. En cómputos globales, 1 de cada 10 demencias es frontotemporal. A pesar de ello, es difícilmente diagnosticable ya que sus síntomas son muy similares a otras demencias, especialmente a la EA. Podemos distinguir distintos tipos de DF [26] [27]:

- Demencia frontotemporal variante conductual: predominan las alteraciones comportamentales entre las que se encuentran cambios en la personalidad, conductas inapropiadas en público, pérdida de empatía, impulsividad, comportamientos repetitivos o compulsivos, apatía y cambios en la dieta.
- Afasias primarias progresivas: predominan las alteraciones del lenguaje por lo que se caracterizan por problemas para expresarse (disartria), para entender el lenguaje (afasia), para leer o para escribir.

Fisiopatología

La DF se caracteriza por una degeneración progresiva de los lóbulos frontales y temporales del cerebro. Macroscópicamente esta degeneración puede llegar a ser tan severa que los giros obtengan un aspecto llamado 'filo de navaja o cuchillo'. A esto se une la gliosis cortical y subcortical, la pérdida cortical de neuronas y la microvacuolación. Microscópicamente se

caracteriza por acumulaciones de proteína TAU (cuerpos de Pick) y depósitos intranucleares o citoplasmáticos de ubiquitinas, así como nudos neurofibrilares [28].

1.2.4 Demencia con cuerpos de Lewy (DCL)

La DCL es la segunda causa de demencia a partir de los 60 años y corresponde al 30% de todas las demencias a nivel mundial. Es difícilmente diagnosticable por lo que se tiende a confundirla con otro tipo de demencias [29].

Como principales síntomas diferenciales tenemos los cambios visuoperceptuales (alucinaciones visuales) y las fluctuaciones en el estado mental que dan lugar a días donde el paciente está muy confuso y otros completamente lúcido. Junto a estos, se pueden observar cambios mentales, parkinsonismo (simétrico y postural), hipotensión ortostática, crisis sincopales, pérdida de conciencia y disfunción ejecutiva (caídas frecuentes, lentitud en los movimientos, etc.) y atencional (lentitud en la atención y concentración) [30].

Fisiopatología

Macroscópicamente, observamos una atrofia cerebral importante con predilección por los lóbulos temporales. A pesar de ello, el diagnóstico se realiza post-mortem y se basa en la observación de cuerpos de Lewy en las neuronas del lóbulo frontal, lóbulo temporal anterior, circunvolución del cíngulo e ínsula [30] [29].

La pérdida de neuronas colinérgicas parece tener relación con los cambios cognoscitivos de la enfermedad, al igual que en la Enfermedad de Alzheimer, mientras, por otro lado, la pérdida de neuronas dopaminérgicas se relaciona con los cambios extrapiramidales que aparecen en la enfermedad de Parkinson [30].

1.3 Neuroanatomía de la visión en enfermedades neurodegenerativas

Multitud de investigaciones afirman que el estudio de métricas de seguimiento ocular no solo mide características básicas del aparato oculomotor, sino que se relacionan también con funciones cognitivas complejas. Esto se debe al intrincado sistema de uniones neurológicas entre gran cantidad de áreas cerebrales y los músculos oculares.

Los movimientos motores voluntarios creados por las regiones corticales se transmiten a estructuras intermedias en el tronco encefálico. Estas estructuras luego coordinan y

transmiten estos planes a los núcleos oculomotores, situados en el mesencéfalo y en la protuberancia, que directamente inervan los músculos extraoculares [1].

Existen múltiples áreas, entre las que se encuentran la corteza fronto-insular, la corteza cingulada anterior, el área motora suplementaria, el coliculi superior y el tálamo, implicadas en tareas de fijación y seguimiento visual. Por otro lado, la corteza prefrontal dorsolateral bilateral se relaciona con la duración de fijación. Mientras la función ejecutiva (sacadas, búsqueda visual y fluida y tareas de cognición social) involucra múltiples regiones corticales y subcorticales: la corteza media está implicada en la activación e inhibición sacádica y los ganglios basales permiten la iniciación del movimiento sacádico mediante la eliminación de la inhibición tónica al colículo superior [17] [31].

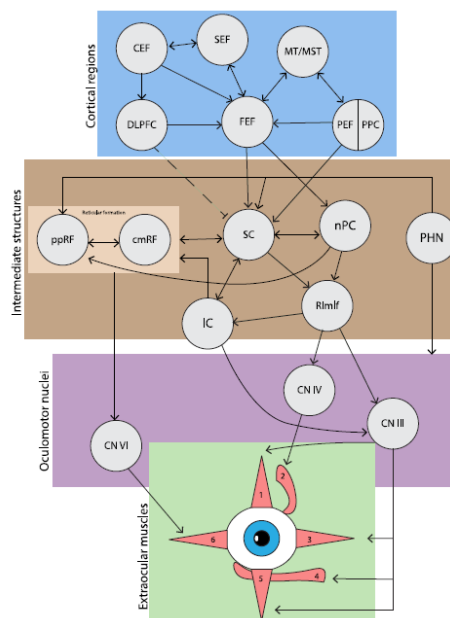


Figura 5. La red oculomotora. Las regiones corticales son las siguientes: campos oculares frontales (FEF), corteza prefrontal dorsolateral (DLPFC), campos oculares cingulados (CEF), campos oculares suplementarios (SEF), lóbulo temporal medio (MT), lóbulo temporal superior medio (MST), campo ocular parietal (PEF) y corteza parietal posterior (PPC). Las regiones intermedias son las siguientes: Colliculus superior (SC), núcleo de la comisura posterior (nPC), núcleo hipogloso prepositus (PHN), núcleo intersticial rostral del fascículo longitudinal medio (RImlf), núcleo intersticial de Cajal (IC), formación reticular pontino-paramediana (ppRF), y formación reticular mesencefálica central (cmRF). Núcleos oculomotores: nervio oculomotor (CN III), nervio troclear (CN IV), nervio abduceno (CN VI). Músculos extraoculares: (1) Recto superior, (2) oblicuo superior, (3) recto medio, (4) oblicuo inferior, (5) recto inferior y (6) recto lateral [1].

Observando la tabla 1, se puede concluir que las regiones implicadas en las enfermedades neurodegenerativas mencionadas se extienden por prácticamente la totalidad de las zonas cerebrales, tanto a nivel superior (telencéfalo) como de estructuras intermedias (tronco encefálico).

Tabla 1. Zonas afectadas por pérdida o disfunción neuronal en las enfermedades neurodegenerativas más prevalentes.

Enfermedad de Alzheimer	Enfermedad de Parkinson	Demencia Frontotemporal	Demencia de Cuerpos de Lewy
Lóbulo frontal		Lóbulo frontal	Lóbulo frontal
Lóbulo temporal	Ganglios basales (mesencéfalo):	Lóbulo temporal	Lóbulo temporal anterior
Lóbulo parietal	Sustancia nigra pars compacta		Circunvolución del cíngulo
	Núcleo estriado		
Lóbulo occipital	Sustancia nigra pars reticulata		Ínsula
Asta temporal del ventrículo lateral	Globus pallidus lateral		
	Globus pallidus medial		
Núcleo estriado piramidal (mesencéfalo)	Núcleo subtalámico		

Comparando con las regiones que conforman el sistema oculomotor podemos observar cómo hay una superposición de regiones con aquellas relacionadas con neurodegeneración, tanto en cercanía como en aferencias neuronales entre áreas:

- Lóbulo frontal: campos oculares frontales (FEF), corteza prefrontal dorsolateral (DLPFC), campos oculares cingulados (CEF), campos oculares suplementarios (SEF).
- Lóbulo parietal: campo ocular parietal (PEF) y corteza parietal posterior (PPC)
- Lóbulo temporal: lóbulo temporal medio (MT), lóbulo temporal superior medio (MST).
- Tronco encefálico:
 - Núcleo: Colliculus superior (SC), núcleo de la comisura posterior (nPC), núcleo hipogloso prepositus (PHN), núcleo intersticial rostral del fascículo longitudinal medio (RImlf), núcleo intersticial de Cajal (IC), formación reticular pontino-paramediana (ppRF), y formación reticular mesencefálica central (cmRF).
 - Protuberancia: nervio troclear (CN IV), nervio abduceno (CN VI).
 - Mesencéfalo: nervio oculomotor (CN III).

Múltiples estudios han demostrado que la información obtenida de los experimentos de seguimiento ocular se correlaciona bien con las escalas de evaluación cognitiva tradicionales [9] [17] [32] [33]. Esto sugiere que el sistema se podría utilizar para evaluar y monitorear los

estados cognitivos y la gravedad y la progresión de la enfermedad en las enfermedades neurológicas, así como su uso en sistemas de recomendación de diagnóstico temprano. Especialmente el uso de evaluaciones oculomotoras, en particular de sacadas verticales (velocidad y ganancia) y antisacadas (porcentaje de acierto), pueden ser clave en el diagnóstico diferencial de síndromes clínicos de demencia.

1.3.1 Movimientos oculares en Alzheimer

Estudios en tareas sacádicas han observado un aumento del número de sacadas, movimientos sacádicos menos precisos, dificultad en el inicio de la sacada y un tiempo de latencia mayor. Respecto a las prosacadas, también se ha observado una disminución en la precisión de visualización de un objeto móvil. En cuanto a sacadas complejas (por ejemplo: predicción del objetivo, supresión de la sacada y decisiones espaciales) también se da una degeneración, produciéndose, por ejemplo, un comportamiento de ignorar al objetivo. Finalmente, respecto al último movimiento sacádico, las antisacadas tiene un rendimiento deteriorado dándose una menor tasa de movimientos correctos y una tasa de errores corregidos muy baja. Toda la degeneración se produce a mayor nivel en el plano vertical respecto al horizontal.

Dado que la corteza prefrontal está implicada en el desencadenamiento de sacadas predictivas y la supresión de antisacadas erróneas, y la corteza parietal posterior puede ser necesaria para las decisiones espaciales, el rendimiento deteriorado en sacadas complejas puede reflejar los profundos cambios neurodegenerativos corticales comúnmente encontrados en estas áreas en la enfermedad de Alzheimer y Parkinson. Por otro lado, el aumento de la latencia se puede correlacionar con la atrofia en regiones corticales visuales, cuyo volumen se encuentra disminuido en EA, y la degeneración en las antisacadas con el volumen disminuido de la materia gris del lóbulo frontal.

Las tareas sacádicas pueden no ser adecuadas como descriptivas de la EA ya que dependen en gran medida de las preferencias visuales personales del paciente.

En cuanto a fijaciones, los pacientes con EA requieren una mayor cantidad de tiempo para fijar el objetivo y estas tienen una duración y frecuencia de fijación más corta.

Por otro lado, en las tareas de seguimiento ocular suave, pasan menos tiempo persiguiendo el objetivo y tienen una tasa de error mayor junto con una aceleración inicial inferior en todas las direcciones, lo que sugiere un déficit de procesamiento visual complejo debido a una perturbación cortical y subcortical. La degeneración en el seguimiento es más prominente a nivel vertical que horizontal, al igual que en las sacadas.

Los pacientes con EA muestran menor curiosidad visual presentando una disminución de la atención visual selectiva hacia estímulos nuevos.

Finalmente, se ha observado que la función visual-espacial se encuentra degenerada dándose, por ejemplo, una precisión espacial disminuida. Esta función está relacionada con el volumen cerebral (lóbulos occipitales y parietales disminuidos) y las regiones corticales extra-estriadas [17] [31] [34] [35].

1.3.2 Movimientos oculares en la Enfermedad de Parkinson

La disfunción ejecutiva en esta enfermedad da lugar a anomalías oculomotoras durante las tareas sacádicas entre las que destacan latencias prosaccádicas más altas, tasas de error antisaccádicas mayores y mayor número de desinhibiciones en las tareas antisaccádicas retrasadas.

La EP también afecta a las redes de ganglios premotor, prefrontal, motor y basal, lo que provoca déficits en el control de la atención. En particular, los pacientes con EP inhiben menos los movimientos sacádicos automáticos frente a un estímulo repentino.

Se produce también un control oculomotor deficiente que se expresa con un control de arriba hacia abajo debilitado. Las métricas de seguimiento ocular se correlacionan con la gravedad de la enfermedad debido a su distonía extraocular característica, lo que puede servir para realizar una evaluación cognitiva en la progresión de la enfermedad en pacientes en todas las etapas de la EP. Esto se demuestra en pacientes recién diagnosticados y no medicados donde la tasa de error antisaccádico es mayor, lo que puede ser un signo temprano de deterioro cognitivo en EP.

Respecto al seguimiento ocular suave hay estudios contradictorios. En algunos no se observa degeneración ni en el plano vertical ni en el horizontal mientras que en otros se concluye que existe un déficit [17] [31] [34] [36].

1.3.3 Movimientos oculares en la Demencia Frontotemporal

En la demencia frontotemporal, se produce una desaceleración psicomotora general que se evidencia tanto en la búsqueda sacádica horizontal voluntaria como en el seguimiento ocular suave horizontal y vertical. Así pues, se muestran latencias y tiempos de reacción aumentados durante la realización de tareas y, por lo tanto, una reducción de las ganancias medias y máximas y de la aceleración inicial [31]. Esto da lugar a, por ejemplo, sacadas hipométricas menores, donde hay una menor ganancia en la primera sacada hacia objetos que aparezcan 5º arriba o abajo o 10º a la izquierda o derecha del punto de fijación central, y mayor tasa de errores en antisacadas. Esta degeneración en el acierto de antisacadas se relaciona con el volumen disminuido de la materia gris del lóbulo frontal presente en la enfermedad.

También tiene lugar una interrupción de la relevancia mientras se mantienen trayectorias de exploración y duraciones de miradas hacia otras áreas de interés cualitativamente normales. De esto se deduce que las señales sociales se detectan, pero no se procesan ni interpretan correctamente [31] [34] [35].

Tabla 2. Resumen de las anomalías oculomotoras por grupo de pacientes observadas en el estudio U. P. Mosimann, R. M. Müri et al, «Saccadic eye movement changes in Parkinson's disease dementia and dementia with Lewy bodies.» [31]. NS = no significativamente distinto de los controles; ↓ = disminución media comparado con los controles; ↓↓ = disminución moderada-severa; ↑ = aumento medio; ↑↑ = aumento moderado-severo. FTD: Demencia frontotemporal; SD: demencia semántica; PA: Afasia no fluente progresiva; AD: Enfermedad de Alzheimer; CBS: Síndrome Corticobasal; PSP: Parálisis Progresiva Supranuclear.

	Sacadas horizontales	Sacadas verticales	Seguimiento horizontal	Seguimiento vertical	Respuesta correcta de antisacadas	Corrección espontánea del error de antisacadas
FTD	↓ Primera ganancia	NS	↓↓ Ganancia ↓↓ Aceleración	↓ Aceleración	↓↓	NS
SD	NS	NS	NS	NS	NS	NS
PA	NS	NS	NS	NS	↓↓	NS
AD	↑↑ Latencia ↓ Primera ganancia	↑↑ Latencia	↓↓ Ganancia ↓↓ Aceleración	↑ Latencia ↓↓ Ganancia ↓↓ Aceleración	↓↓	↓↓
CBS	↑ Latencia ↓ Primera ganancia	↑↑ Latencia ↓↓ Ganancia final	↓↓ Ganancia ↓↓ Aceleración	↑ Latencia ↓ Aceleración	↓↓	↓↓
PSP	↓ Primera ganancia ↓↓ Velocidad	↑↑ Latencia ↓↓ Ganancia ↓↓ Velocidad	↑↑ Latencia ↓↓ Ganancia ↓↓ Velocidad	↑↑ Latencia ↓ Ganancia ↓↓ Velocidad	↓↓	↓↓

1.3.4 Movimientos oculares en Demencia por Cuerpos de Lewy

Las deficiencias que presenta la DCL son muy similares a las que se observan en la EP. Así pues, existe deterioro en la ejecución de sacadas reflexivas y en el rendimiento de sacadas complejas (por ejemplo: predicción de objetivo, supresión de sacada y decisiones espaciales) viéndose aumentada su tasa de error y su latencia. También encontramos alteraciones en los espacios entre sacadas, solapamiento entre ellas y aumento de la latencia en las antisacadas. Todo esto es debido a la coexistencia de neurodegeneración en áreas corticales y subcorticales. Por otro lado, la ejecución de las sacadas reflexivas es un elemento diferenciador respecto a la EA, enfermedad a la que se asemeja altamente, ya que en esta existe un deterioro predominantemente en sacadas complejas debido a la existencia, principalmente, de neurodegeneración en la zona cortical. Así pues, encontramos latencias más largas en DCL que en EA y una intensidad de la primera sacada menor. Las sacadas reflexivas son desencadenadas por el campo ocular parietal y las complejas por el campo ocular frontal, teniendo ambas conexiones excitatorias con el coliculus superior y el estriado, y recibiendo retroalimentación de los ganglios basales a través de los tractos talamocorticales. La combinación de la degeneración ejecutiva de las sacadas y de su control sugiere una disrupción en el procesamiento de las señales en estructuras subcorticales [31].

Respecto a la percepción visual en la EP y la DCL, existen deficiencias en la discriminación visual, movimiento espacial y reconocimiento de formas de objetos predominando los déficits en orientación, trazabilidad y lectura de nombre de colores en la DCL [25]. Finalmente, también se ha encontrado en la DCL un aumento en la omisión de objetos visuales durante el seguimiento ocular.

2

Diseño y desarrollo

2.1 Fase 1: Elección de los dispositivos

Durante esta fase se realizó un inventario de todos los elementos necesarios para la obtención de los datos de movimiento ocular susceptibles de ser utilizados en herramientas de recomendación de diagnóstico de enfermedades neurológicas. Los principales son el ordenador y el propio dispositivo de seguimiento ocular.

La elección del ordenador se basó en las existencias disponibles en el laboratorio ya que para la realización del proceso completo no se requiere especificaciones complejas o de alta gama. Así pues, el ordenador a utilizar fue un PC con procesador i7, 16GB de RAM y sistema operativo Windows.

Junto con él se escogió una pantalla Asus de 23.8 pulgadas siguiendo los requerimientos del fabricante del eye-tracker para optimizar la obtención de los datos [37].

A la hora de escoger el sistema de medición, se tuvo en cuenta las distintas opciones existentes en el mercado junto con las dos opciones disponibles en el laboratorio, un dispositivo Tobii modelo X2-30 y otro de la compañía EyeTribe.

En primer lugar, hubo que decidir qué tipo de dispositivo escoger dentro de los existentes en la industria del seguimiento ocular. Estos tienen distintos tipos de prestaciones: inalámbricos, con estabilización de cabeza, móviles (gafas) y sistemas integrados/incorporados. El ámbito de uso de cada uno está bastante restringido en base a sus especificaciones y limitaciones (ver Tabla 3).

Tabla 3. Comparativa de los dispositivos de seguimiento ocular actualmente disponibles en el mercado [38].

Tipo de sistema de Eye-tracking	Uso habitual	Limitaciones
Inalámbrico	<ul style="list-style-type: none"> • Experimentos en psicología, neurociencia y visión basados en una pantalla. • Participantes en edad infantil o neuro-comprometidos. • Comunicación asistida. • Usabilidad web y de software. • Estudios de mercado de anuncios televisivos, páginas web y publicidad. • Experimentos con EEG, fNIRS o biosensores. • Juegos online. 	<ul style="list-style-type: none"> • Los participantes se tienen que mantener dentro del rango de la cámara. • Movimiento ocular se graba en 1 plano (generalmente en un monitor de TV). • El movimiento de cabeza excesivo puede dar lugar a imprecisión y desconexiones durante el rastreo ocular. • No funciona bien con luz solar.
Móvil	<ul style="list-style-type: none"> • Experimentos sobre el comportamiento humano en el mundo real. • Entrenamientos deportivos, rehabilitación, biomecánica, comunicaciones, factores humanos, quinesiología, ergonomía. • Punto de venta, marketing de estanterías en tiendas. • Vehículos y simuladores de alta fidelidad. 	<ul style="list-style-type: none"> • Los datos son relativos al campo de visión o a la escena del video, no a un sistema de coordenadas absolutas fijas. • Los datos estadísticos necesitan un análisis más subjetivo. • Mayor variabilidad en resultados debido a la mayor libertad de los participantes. • Movimientos oculares excéntricos no rastreados de forma adecuada.
Con estabilización de cabeza	<ul style="list-style-type: none"> • Experimentos de alta fidelidad en neurociencia, fisiología, visión o psicología. • Experimentos donde la tasa de muestreo y la precisión sean más importantes que la comodidad del participante. 	<ul style="list-style-type: none"> • Incómodo para experimentos de larga duración. • No muy portátil. • Los participantes no se encuentran en su entorno natural. • Inviabile para participantes de ciertas poblaciones.
Integrados	<ul style="list-style-type: none"> • Realidad aumentada (RA) • Realidad virtual (VR). • Dispositivos en tiempo real como cámaras. • Análisis médicos y herramientas quirúrgicas oculares. 	<ul style="list-style-type: none"> • Útil solo para el dispositivo en cuestión. • En RA/VR, los pacientes tienen que tolerar el dispositivo. • Habitualmente altamente especializado para un uso en particular.

Teniendo en cuenta lo anterior, se decidió utilizar un dispositivo remoto ya que es el más comúnmente usado para experimentación en personas neuro-comprometidas y/o en seguimientos de larga duración. Estos dos factores son claves en nuestro proyecto. Además, este dispositivo no requiere contacto directo con el paciente lo que favorece una interacción más natural con el ordenador sin alterar el comportamiento de la persona durante la lectura. Finalmente, otro punto positivo a largo plazo sería la posibilidad de integrar el seguimiento ocular con otras tecnologías en investigación neurológica como la EEG, lo que nos aportaría mayor cantidad de datos para mejorar el sistema de diagnóstico de las enfermedades neurodegenerativas [38].

Respecto a las limitaciones observadas, el hecho de mantener al paciente en un área restringida cercana al dispositivo no supone ningún problema ya que la actividad a realizar es estática y situada en una pantalla (un único plano). Esto también favorece el hecho de evitar movimientos de la cabeza durante la lectura. Aun así, a cada voluntario se le insta a evitar mover la cabeza durante el desarrollo de la prueba y a mantener los ojos fijos en el monitor. Junto con esto, para mantener el área de trabajo del dispositivo limpia de otras interferencias, será la persona que lea en ese momento la única situada en el rango de actuación del *eye-tracker* evitándose así que el dispositivo realice seguimientos cruzados

entre varias personas al mismo tiempo. Por último, el reflejo de la luz del sol en los ojos del participante es fácilmente controlable mediante el montaje del sistema en un ambiente de luz artificial lejos de ventanas, por lo que los registros no se realizaron frente a elementos como ventanas o fuentes de luz directa. Como medida preventiva ante la posible influencia de variaciones lumínicas en los valores de medición, se empleó un luxómetro para medir los luxes de luz artificial existentes en la sala en cada prueba [38].

A continuación, se compararon las características, así como la relación calidad precio de los dispositivos disponibles en el laboratorio junto con los comercializados por las 5 mejores compañías en equipos de seguimiento ocular, Tobii, SMI, EyeLink, Smart Eye y LC Technologies-Eyegaze (Tabla 4).

Observando las características arriba mencionadas junto con otros datos a tener en cuenta como la portabilidad (tamaño), el software de adquisición y análisis de datos (incluido o no), la compatibilidad con diversos sistemas operativos y la técnica usada para el reconocimiento ocular y el cálculo de la posición de mirada, entre otras, concluimos que el dispositivo Tobii que tenemos a disposición en el grupo cumple ampliamente los requisitos para realizar un proyecto de alta calidad. Además, el uso de un dispositivo Tobii viene respaldado por la trayectoria de dicha marca siendo esta la indiscutible líder en el mercado.

En una futura posible ampliación respecto a los centros involucrados en el estudio se podría plantear el uso de la herramienta de diagnóstico con los dispositivos Gazepoint GP3 y Tobii Pro Nano debido a que ambos tienen unas magníficas especificaciones junto con un precio asequible para la mayoría de instituciones.

Tabla 4. Comparación de las características técnicas de los diversos dispositivos para eye-tracking existentes en el mercado y dirigidos al ámbito de la investigación de las 5 empresas líderes del área.

Nombre	Velocidad de muestreo	Precisión	Exactitud	Calibración	Rango de distancia	Tamaño de pantalla	Número de cámaras	Precio aproximado
Gazeoint GP3	60 y 150 Hz	0.5-1°		5 y 9	50-80 cm	Hasta 24"	2	1.995 €
Eyegaze Edge	50 Hz	0.45°	0.2°	5, 9 y 13	43-73 cm	Tablet incluida 12.5"	2	-
Smart Aurora	60 y 120 Hz	-	0.3°	4 puntos	50-80 cm	-	-	5.670 y 8.010 €
Smart Eye AI-X	60 Hz	0.1°	0.5°	-	45-85 cm	Hasta 24"	2	2.790 €
Eyelink Portable Duo	250, 500, 1000 y 2000 Hz	0.01°	0.25-5°	13 puntos	42-62 cm	Depende de la distancia	2	-
Eyelink 1000 Plus	250, 500, 1000 y 2000 Hz	0.25°	0.25-0.5°	9 puntos	40-70 cm	Depende de la distancia	2	-
SMI Livetrack Lightning	Desde 60 hasta 500 Hz	0.05°	0.25-0.5°	-	25 cm	Depende de la distancia	2	6.900 – 15.900 €
Tobii Pro Spectrum	60, 120, 150, 300, 600 or 1200 Hz	0.01°	0.3°	5 puntos	55-75 cm	Óptimo 24"	2	21.900 -31.900 €
Tobii Pro Fusion	Hasta 250 Hz	0.04°	0.3°	5 puntos	50-80 cm	Óptimo 24"	2	12.900 – 16.900 €
Tobii Pro Nano	60 Hz	0.1°	0.3°	5 puntos	45-85 cm	Hasta 19"	1	3.200 €
Tobii X2-30	30 Hz	0.4°	0.32°	5 puntos	40-90 cm	Hatsa 25"	2	20.000 €
EyeTribе	30 y 60 Hz	0.5°	0.5°	5,9 y 12	45-75 cm	Desde 24"	-	99 \$

2.2 Fase 2: Diseño del estudio

En esta fase se siguieron las pautas reflejadas en los distintos manuales proporcionados por el fabricante del dispositivo, tanto a nivel hardware (posicionamiento) como software (calibración y protocolo de obtención de datos).

El primer paso en el desarrollo del experimento con cada voluntario es la disposición espacial de los instrumentos necesarios. El sistema entero se situará mirando a una pared, alejado de las ventanas, para evitar interferencias por la luz solar.

El monitor se colocará en una mesa de tal forma que el centro de la pantalla se sitúe ligeramente por debajo de la altura de los ojos (ver Figura 5). Se medirá el centro de dicho monitor y se colocará el dispositivo de seguimiento ocular de tal forma que el centro de este coincida con el centro de la pantalla en la parte inferior ya que, por la forma de la pupila, se obtendrá una imagen de esta más clara que situándolo en la parte superior del monitor. A continuación, se configurarán los parámetros espaciales de la disposición.

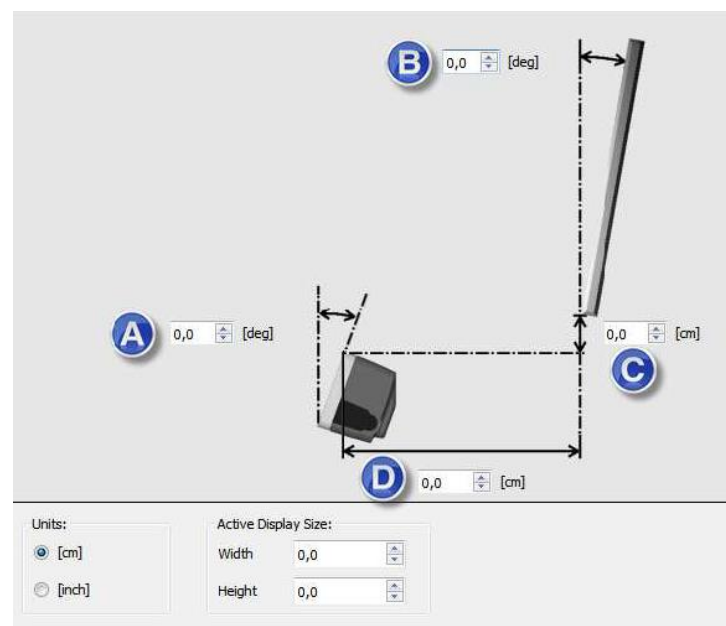


Figura 5. Parámetros de configuración espacial entre el monitor y el dispositivo de seguimiento ocular. Unidades: unidad elegida para medir distancias; Tamaño del área de visualización activa: área donde se lleva a cabo la visualización de la calibración. Debe ser un 25% mayor que el tamaño real del monitor. A: ángulo del dispositivo de seguimiento ocular; B: ángulo entre el plano vertical y el monitor. En caso de estar inclinada hacia delante se pondría un signo negativo; C: Altura entre la última fila de píxeles del monitor y el punto de medida del dispositivo; D: Distancia entre el punto de medida del dispositivo y el monitor.

Se escogieron centímetros como unidad de medida y se midieron los demás parámetros obteniéndose un área de visualización de 56.2cmx20cm, un ángulo del dispositivo de 37° (el recomendado para una precisión óptima con un monitor de 24" y un relación de aspecto de 16:9), un ángulo entre el plano vertical y el monitor de 0° ya que el monitor es una pantalla

recta y una altura y una distancia entre el monitor y el dispositivo de 0 cm ya que el dispositivo está pegado a la parte inferior de la pantalla.

A continuación, el voluntario se sentará en una silla cómoda, pero sin ruedas, y se colocará a una distancia de 55 cm del dispositivo. A pesar de que para este sistema la distancia recomendada es de 65 cm, pudimos observar tras múltiples calibraciones que, comparativamente, los mejores resultados se obtenían acortando la distancia un poco hasta los 55 cm con un margen de ± 2 cm. En caso de que se tengan problemas con la calibración y se tenga que repetir, el resultado mejorará moviendo al paciente ligeramente hacia delante o hacia atrás. Esta disminución en la distancia recomendada afectaría a la libertad de movimiento de cabeza permitida por el aparato (50x36 cm) ya que esta se asegura a una distancia de 70 cm, pero la influencia de este factor es prácticamente nula en los resultados ya que durante la lectura del texto se insta profundamente a los pacientes a mover únicamente los ojos.

Una vez el paciente está situado enfrente del monitor, se medirá la intensidad lumínica de la sala con un luxómetro para analizarla como posible factor en caso de observar posteriormente diferencias significativas entre voluntarios. Para ello se colocará en la pantalla un texto con los mismos colores y la misma estructura que el que se va a utilizar en la prueba real para igualar las condiciones ambientales. Tanto para los dos textos utilizados en el experimento como para el de prueba para medir los lúmenes se ha utilizado el mismo tipo de letra (Arial), colores (texto negro y fondo blanco), tamaño de fuente (22) e interlineado (2.0). La uniformidad de los textos facilita el análisis y aumenta la precisión de este ya que las coordenadas de las líneas son las mismas.

Consecutivamente, se iniciará la calibración y validación. Esta parte es la más crítica de todo el proceso ya que se basa en la obtención de datos específicos de cada individuo previamente a la realización de su prueba para ajustar el modelo en 3D del ojo a sus características propias e individuales y asignar en cada instante a la posición de la mirada las coordenadas de la pantalla del ordenador precisas (punto de referencia de nuestra grabación, POR).

Se escogió el método de calibración 'Jumping point' en el que una serie de puntos van apareciendo en distintas coordenadas consecutivamente (ver Figura 6). Este método es el más utilizado en los estudios de investigación que utilizan este tipo de dispositivos. Se escogieron 5 puntos de calibración (uno en cada esquina junto con otro central) y cada uno con una duración en pantalla de 0.5 segundos para dar la oportunidad de centrar la mirada al usuario. Este número de puntos se decidió debido a que da lugar a un buen resultado a la vez que disminuye el tiempo de duración de esta fase y la fatiga visual derivada de la persecución de puntos en la pantalla, punto de gran importancia en nuestros grupos diana, personas mayores sanas y con enfermedades neurodegenerativas. En base a la precisión y exactitud que se observe en la imagen final, se repetirá la calibración o se seguirá adelante con la prueba.



Figura 6. Ejemplo de calibración no aceptable para pasar a la prueba. No se observan tanto puntos rojos como azules en todos los puntos lo que implica que alguno de los dos ojos no ha sido captado por el aparato. Se puede observar una baja precisión (dispersión entre los puntos) de los valores en el punto situado abajo a la izquierda aunque la exactitud general (distancia entre el valor real y los obtenidos) es bastante elevada teniendo en cuenta el tamaño total de la pantalla.

La exactitud de cada muestra nos indicará la diferencia entre la posición del foco de mirada en la pantalla y la calculada por el dispositivo.

Por otro lado, la precisión es la capacidad del dispositivo de reproducir correctamente el punto de mira. En nuestro caso, esta métrica se calcula mediante la media cuadrática o RMS de una cantidad de puntos sucesivos para cada uno de los ojos y para la media de los dos.

Una vez calibrado, el modelo generado se evaluará a partir de los datos que se obtengan del dispositivo en la pantalla siguiente realizando el mismo ejercicio. A partir de la validación, obtendremos un valor final que nos indicará lo óptima que ha sido la calibración.

Todo este proceso se realiza gracias a la API del dispositivo Tobii, que es la encargada de realizar el mapeo entre la posición de los ojos y las coordenadas de la pantalla.

Finalmente, se llevará a cabo la prueba de lectura. Hemos escogido una prueba de lectura ya que se relaciona directamente con funcionalidades neurológicas profundas de procesamiento cognitivo. Además, es una actividad familiar y simple para los voluntarios y que se relaciona con diferentes capacidades cognitivas, comprometidas en pacientes con enfermedades neurodegenerativas.

Así pues, se realizará todo el proceso descrito con anterioridad para la lectura de dos tipos de textos cortos por cada individuo con la intención de variar la complejidad de los relatos

y, así, obtener datos en distintas condiciones. El primer texto será el cuento de Caperucita Roja mientras que el segundo será un fragmento de Don Quijote de la Mancha (Anexo I). Ambos textos están formados por 10 líneas y tienen el mismo formato de tipo, color y tamaño de letra.

2.3 Fase 3: Obtención de los datos

Una vez estructurado el experimento y controladas las variables que pueden interferir en los resultados, se procedió a la búsqueda de voluntarios.

Tras contactar con varias residencias de la tercera edad y hospitales se consiguió establecer un convenio con la residencia Domus IV en Alcalá de Henares y con el Hospital de los Madroños para el reclutamiento y participación de sus residentes y pacientes, respectivamente, en el estudio.

Así pues, tras redactar el correspondiente convenio, el consentimiento informado y mandar el protocolo de obtención de datos (Anexo II) se obtuvo un grupo de 15 participantes provenientes de la residencia y 12 del hospital. De los voluntarios de la residencia, 9 forman parte del grupo estudio al estar diagnosticados con algún tipo de deterioro cognitivo y 6 conforman el grupo control. Mientras, los pacientes del hospital se dividen en grupo control (4), personas que han sufrido un accidente cerebrovascular (5) y personas que han sufrido un traumatismo craneoencefálico (3), ambas patologías relacionadas con una pérdida importante en las funciones cognitivas.

Los criterios seguidos para la inclusión o exclusión dentro de los grupos estudio y control en la residencia fueron los siguientes:

- Inclusión
 - Edad entre 70-100 años.
 - Nivel académico: medio-alto (grado medio o superior en estudios).
 - Grupo estudio: Pacientes diagnosticados con deterioro cognitivo (Enfermedad de Alzheimer, Demencia por Cuerpos de Lewy, Demencia por enfermedad de Parkinson, Demencia Frontotemporal, Demencia vascular, etc) independientemente del estadio en el que se encuentren.
- Exclusión:
 - Problemas visuales que impidan la lectura o el funcionamiento correcto del dispositivo: daltonismo, cataratas, ceguera, hipermetropía, astigmatismo o miopía severas no corregidas, estrabismo, etc.
 - Uso de gafas o lentillas que impidan el seguimiento ocular. Se intentará realizar la prueba con estos pacientes y se continuará siempre y cuando la calibración sea adecuada. En caso de no conseguir una precisión y exactitud

adecuadas en varias repeticiones de la calibración dicho paciente se descartará.

- Epilepsia fotosensible: algunas personas diagnosticadas con esta enfermedad son susceptibles a convulsiones epilépticas o pérdida de conciencia cuando se exponen a ciertas luces intermitentes o patrones de luz en la vida cotidiana. Pueden tener problemas tanto con la pantalla donde se muestran los textos como con las fuentes de luz del propio rastreador ocular. Se estima que alrededor del 3-5% de las personas con epilepsia tienen este tipo de epilepsia fotosensible. Muchas personas con epilepsia fotosensible experimentan un aura o sienten sensaciones extrañas antes de que ocurra la convulsión por lo que se explicará a los voluntarios que en caso de percibir síntomas como este alejen los ojos del dispositivo.
- Marcapasos/desfibriladores cardíacos: el aparato de seguimiento ocular contiene imanes. Los campos magnéticos pueden interferir con la función de los marcapasos cardíacos y los desfibriladores cardioversores implantables cuando no se respeta una distancia de mínimo 15cm entre el rastreador y el dispositivo médico. Para evitar riesgos se decidió eliminar a pacientes con este tipo de instrumentos médicos del estudio.
- Grupo control: Enfermedades neurodegenerativas, demencias o cualquier otra que influya en la cognición de la persona (enfermedades psiquiátricas, personas con discapacidad mental, ictus, etc).

Estos mismos criterios se siguieron para los pacientes del hospital a excepción del criterio diagnóstico de inclusión en el grupo estudio que será el diagnóstico médico de haber sufrido algún accidente cerebrovascular o traumatismo craneoencefálico que haya afectado a las funciones cognitivas.

Tras el estudio de la lista de participantes proporcionada por ambos centros, se llevó a cabo una sesión de instrucción a las personas responsables de la realización de las pruebas. En ella se explicó punto por punto los pasos a seguir, se hizo algunas pruebas in situ y se mostraron ejemplos de validaciones adecuadas y otras que no pasarían el corte y se deberían repetir.

A lo largo de todo el periodo de recogida de datos hubo un contacto continuo entre los centros y nosotros para la resolución de dudas, especialmente en cuanto a las validaciones correctas o no y la consecuente repetición de la prueba a ciertos voluntarios, y para asegurar que los protocolos se estaban siguiendo adecuadamente.

Todos los datos obtenidos fueron tratados de forma anónima llegándonos a nosotros únicamente un sistema de identificación creado por el propio centro (iniciales de los nombres y apellidos de cada individuo para los voluntarios de la residencia y un número asociado a cada paciente para los recibidos del hospital) junto con su edad y diagnóstico.

2.4 Fase 4: Elección de los métodos de análisis de datos

Con los datos de cada paciente ya en nuestras manos el primer paso fue el tratamiento previo de los datos, eliminando los artefactos con valores -1 (pérdida de conexión entre los ojos del voluntario y el dispositivo) y estructurándolos de tal forma que se puedan introducir en los algoritmos escogidos.

Debido a la escasa participación de voluntarios que conseguimos en el corto plazo de tiempo disponible en el contexto de la pandemia, decidimos dividir cada lectura de cada paciente en cuatro y, así, aumentar las series temporales tanto en el entrenamiento para mejorar la clasificación como en la validación para aumentar la robustez de los modelos. La división se realizó de forma equitativa, de tal forma que de cada lectura se obtuvieron cuatro fracciones con la misma cantidad de mediciones. De estas cuatro partes, se escogió de forma aleatoria una para la validación posterior del modelo y las otras tres para su entrenamiento. El hecho de separar sin criterio una de las muestras para el testeo y el resto para el entrenamiento aumenta la robustez del modelo ya que en caso de que exista variabilidad en una región en concreto habrá muestras de dicha región tanto en la parte test como en la de entrenamiento, adaptándose mejor a cualquier variación existente.

A continuación, estos datos se presentan a los algoritmos escogidos para obtener el resultado final de precisión en la clasificación. En este punto, los datos se dividieron para realizar dos tipos de estudios de clasificación distintos:

- Estudio inicial: los datos provenientes de la residencia se utilizaron para investigar la posibilidad de apoyo al diagnóstico de neurodegeneración. En este caso, no tendremos en cuenta el tipo de enfermedad que sufre el voluntario y solo existirán dos clases (control y estudio). En este paso es donde compararemos los algoritmos escogidos para, en la fase siguiente, analizar los datos provenientes del otro centro con el algoritmo más eficiente en la clasificación.
- Estudio secundario: aquí utilizaremos los datos provenientes del hospital para investigar la posibilidad de no solo apoyar un diagnóstico de neurodegeneración, si no también de otros tipos de enfermedades neurológicas. Además, se extenderá esta clasificación no solo a obtener un resultado diagnóstico genérico, como puede ser el de neurodegeneración, si no a conseguir diferenciar las diversas enfermedades entre sí y obtener un diagnóstico preciso del trastorno que afecta al individuo.

En la literatura se han propuesto numerosos algoritmos para tratar los datos del seguimiento ocular, incluyendo distintas técnicas de selección de características [39] [40] [41] [42] [43]. Para escoger los algoritmos de aprendizaje automático a utilizar se buscaron aquellos que pudieran trabajar con series temporales multivariantes ya que, el principal objetivo de este estudio es observar si utilizando las series temporales del seguimiento ocular y, en concreto, las relaciones existentes entre datos cercanos temporalmente se puede conseguir un sistema de clasificación robusto para la asistencia en el diagnóstico. A pesar de que ya existen estudios de movimientos oculares en enfermedades neurológicas [8] [9] [17] [31]

[34] [35], en la mayoría se utilizan medidas promediando o descartando la estructura temporal de los datos, sin establecer relaciones entre ellos.

Así pues, se decidió el uso de dos algoritmos que permiten considerar la estructura secuencial de los datos: red neuronal Long Short Term Memory (LSTM) y Multivariate Time Series Classification with Learned Discretization (SMTS), para establecer posteriormente comparativas entre ambos y decidirnos, finalmente, por el que mejor resultado nos proporcione.

2.4.1 Long Short-Term Memory (LSTM)

En primer lugar, se escogió una red neuronal recurrente debido a su funcionamiento ya que se basa en la transformación del input mediante una función (W_y y W_x) pasando a través de un estado escondido (h) que se actualiza (W_h) usando esta información en el procesamiento de la siguiente entrada. De esta forma la capa oculta captura la relación entre datos vecinos y así se mantiene la estructura temporal [44] [45] [46].

La diferencia de este tipo de red con el resto de redes neuronales es que cada entrada llega a una celda de la capa oculta formada por tres puertas: *input*, *output* y *forget*. Estas puertas controlan el flujo de información para predecir el output de tal forma que solo retiene la relevante y olvida la que no lo es. De esa forma se evita dependencias no deseadas a largo plazo, es decir, que no se capte una relación de influencia entre varios datos por estar estos situados temporalmente lejos [44] [45] [46].

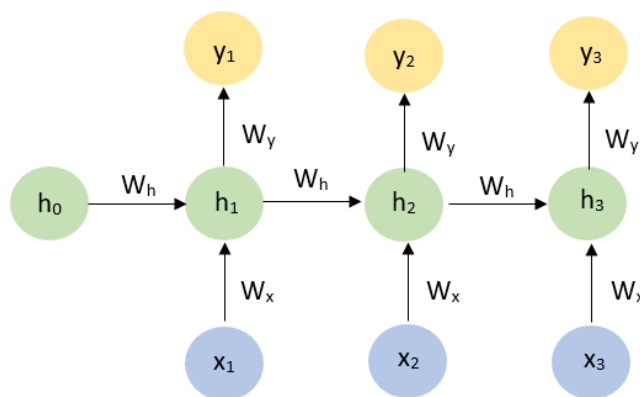


Figura 7. Estructura funcional de una Red Neuronal Recurrente. x_1 , x_2 , x_3 : inputs; w_y , w_x y w_h : pesos de las funciones de transformación de los datos; h_0 , h_1 , h_2 , h_3 : valores para cada input en la capa oculta; y_1 , y_2 , y_3 : output [44].

Los pesos (W_y , W_x y W_h) se van ajustando a medida que la red aprende a través de un algoritmo de retropropagación a través del tiempo (BPTT), específico de datos de secuencia, que determina los gradientes de la función de error. Este sistema se basa en el cálculo de errores y su suma en cada paso del tiempo para ajustar los parámetros del modelo [47] [46].

El algoritmo BPTT tiene dos problemas: los gradientes explosivos y desvanecientes. El primero se da cuando el gradiente (derivada de la función de pérdida) es demasiado grande dando lugar, eventualmente, a pesos en el modelo demasiado grandes. El segundo se produce cuando los gradientes son demasiado pequeños de tal forma que los pesos tienden a 0 y el algoritmo deja de aprender. Ambos problemas se evitan al usar LSTM ya que este tiene únicamente una capa escondida, lo que simplifica la red [47] [46].

Por otro lado, nuestros datos son series temporales por lo que la entrada serán sus múltiples valores, mientras que la salida queremos que sea una etiqueta: paciente sano o enfermo. De esta manera, nos encontramos ante una red neuronal many-to-one donde la capa de salida estará formada por una única neurona que nos dará un único output, 0 para sano y 1 para enfermo.

Para obtener el mejor modelo de entrenamiento utilizando este algoritmo, se pueden variar los valores de los distintos parámetros que definen a la red [48] [49] [50]:

- *Units*: Dimensión de cada celda de la capa oculta.
- *Dropout_rate*: Probabilidad de cambiar el valor que llega a una capa procedente de la anterior a 0 de tal forma que ese valor se elimina y no se tiene en cuenta para el entrenamiento de la red. Es una forma de disminuir el posible sobreajuste que se puede dar en las redes neuronales creando un modelo más robusto.
- *Activation*: Función de activación de las neuronas.
- *Batch_size*: Número de elementos que se muestran a la red antes de realizar una actualización en los pesos. Cuanto más pequeño sea el valor, más rápido aprende la red, aunque, también, más inestable es el aprendizaje.
- *Epochs*: Número de veces que se corre el modelo dando lugar a una evaluación final. Un modelo con los mismos parámetros iniciales da lugar a distintos resultados cada vez que se entrena debido a que se establecen en cada vez unas condiciones iniciales aleatorias y, por tanto, distintas. Cuantas más veces se entrene el modelo más exacto será este, pero también aumentará el riesgo de sobreajuste.
- *Optimizer*: Clase con información añadida para mejorar la velocidad y el rendimiento en el entrenamiento del modelo.
- *Loss function*: Función de cálculo del gradiente que se utilizará para actualizar los pesos de la red neuronal. Los valores obtenidos de esta función nos permitirán observar si existe sobreajuste en el modelo final.

2.4.2 Multivariate Time Series Classification with Learned Discretization (SMTS)

Este modelo de clasificación fue desarrollado por Mustafa Gokce Baydogan y George Runger para el análisis de series de datos tanto univariantes como, y en especial, multivariantes.

El funcionamiento se basa en un modelo de Random Forest (RF) con una codificación temporal para tratar la estructura secuencial de los datos. Este proporciona una serie de ventajas para el tratamiento de series temporales. Entre ellas se encuentra el manejo eficiente de espacios de características de gran dimensión, múltiples clases y valores perdidos. Además, se permiten modelos no lineales e interacciones entre características, es invariante a escala y robusto para valores atípicos. Finalmente, este modelo es computacionalmente favorable en caso de grandes conjuntos de datos.

La principal mejora en este algoritmo es que la serie de datos se trata de forma conjunta e individual, incluyéndose todos los atributos para establecer información de las relaciones entre ellos. Lo primero que se lleva a cabo, por lo tanto, es la estructuración de los datos previamente a su introducción en el algoritmo de RF.

Ecuación 1. Estructura de cada serie temporal en el algoritmo SMTS. Cada columna representa un atributo de un total de M atributos y cada fila son los valores de dicho atributo en la serie (n) en un instante de tiempo t.

$$X^n = \begin{bmatrix} x_1^n(t_1) & \cdots & x_M^n(t_1) \\ \vdots & \ddots & \vdots \\ x_1^n(t_t) & \cdots & x_M^n(t_t) \end{bmatrix}$$

Cada columna se agrupa en lo denominado una instancia de tal forma que esta matriz se convierte en un vector. Se agrupan todos los vectores de todas las series en una matriz formando la matriz $D_{NT \times M}$.

Ecuación 2. Matriz donde convergen todas las series temporales cuyas dimensiones vienen definidas por el número de atributos existentes (M) y el número de series temporales (N). Cada valor de cada celda es un vector formado por los valores de un atributo específico en una serie temporal específica a lo largo del tiempo (t).

$$D_{NT \times M} = \begin{bmatrix} x_1^1 & \cdots & x_M^1 \\ \vdots & \ddots & \vdots \\ x_1^N & \cdots & x_M^N \end{bmatrix}$$

Donde la instancia $x_1^1 = [x_1^1(t_1) \cdots x_1^1(t_t)]$

A esta matriz nueva se le aplica una función de mapeo ($\Phi_{NT \times (2M+1)}$) para tener en cuenta las posibles relaciones entre los datos cercanos en el tiempo. Gracias a esta función se crea una columna con el índice temporal y, para cada atributo, una columna cuyo valor es la diferencia entre el dato en dicha posición y el que le precede en el tiempo.

Ecuación 3. Matriz formada por las instancias originales junto con una columna de referencia temporal (vectores de los valores del tiempo desde el tiempo 0 al tiempo t) y las columnas cuyos valores son las diferencias entre el valor original y aquel que lo precede.

$$D_{NT \times (2M+1)} = \begin{bmatrix} T & x_1^1 & x_1^{1*} & \cdots & x_M^1 & x_M^{1*} \\ T & \vdots & \vdots & \ddots & \vdots & \vdots \\ T & x_1^N & x_1^{N*} & \cdots & x_M^N & x_M^{N*} \end{bmatrix}$$

Donde $T = [t_1 \cdots t_t]$ y $x_1^{1*} = [x_1^1(t_1), x_1^1(t_1) - x_1^1(t_0) \cdots x_1^1(t_t), x_1^1(t_t) - x_1^1(t_{t-1})]$

De esta forma, los datos que usamos para el entrenamiento serán los valores de los atributos de la serie, el índice de tiempo y una tercera columna donde se calculará la diferencia de cada dato con el anterior para establecer relaciones entre ellos.

La matriz $D_{NT \times (2M+1)}$ se utilizará como input para el primer modelo RF de este algoritmo. Teniendo en cuenta que estamos usando una matriz de instancias, este Random Forest se denomina RF_{ins} . La precisión final del modelo depende en gran medida de este punto y de los valores de los dos parámetros variables que se pueden modificar en RF_{ins} . Estos parámetros son el número de árboles a usar en el entrenamiento (J_{ins}) y el número de nodos finales (R). Cuanto más altos sean estos valores más exacta será la clasificación, pero tendremos, también, mayor riesgo de sobreajuste. A pesar de que no existe una poda directa de los árboles, el hecho de definir un número de nodos finales (R) menor al número de instancias hace que se prevenga el sobreajuste. Por otro lado, la elección aleatoria en cada árbol del conjunto de instancias de entrenamiento y de las características a tener en cuenta en cada nodo para diferenciarlas reduce la varianza y la carga computacional del modelo. Todo esto da lugar a un algoritmo robusto y con una probabilidad de sobreajuste baja.

Una vez hemos pasado nuestra matriz de datos por el número de árboles fijado (J_{ins}) usaremos un sistema de Bag-Of-Words para crear la matriz de frecuencias relativas para cada serie temporal (X^n). Cada vector de dicha matriz ($H(X^n)$) es una sucesión de los porcentajes de instancias de dicha serie presentes en cada uno de los nodos finales (R) en cada uno de los árboles (g_j) que conforman el RF_{ins} .

Ecuación 4. Matriz para cada serie (X^n) de los porcentajes de instancias pertenecientes a dicha serie en cada nodo final de cada árbol del RF_{ins} . Las dimensiones de esta matriz serán, por lo tanto, el número de nodos finales (R) por el número de árboles utilizados en el RF (J_{ins}).

$$H(X^n) = \begin{bmatrix} \%X^{nr_1}_{g_1} & \cdots & \%X^{nr_R}_{g_1} \\ \vdots & \ddots & \vdots \\ \%X^{nr_1}_{g_{J_{ins}}} & \cdots & \%X^{nr_R}_{g_{J_{ins}}} \end{bmatrix}$$

El conjunto de las matrices de frecuencias relativas para todas las series se utiliza como input para el segundo algoritmo de Random Forest (RF_{ts}) junto con las clases asociadas a cada serie. El uso de RF como segundo algoritmo se elige debido a que las dimensiones de la matriz resultante de RF_{ins} puede llegar a ser grande y RF además de manejar dicha dimensionalidad, es un clasificador escalable que maneja interacciones y correlaciones.

Los parámetros que podemos variar en este modelo para lograr un mejor ajuste son R y J_{ins} . Ambos van a definir tanto el tiempo de entrenamiento como la precisión final del algoritmo. El tiempo de computación disminuye exponencialmente cuanto menor sea R sin una pérdida sustancial en el rendimiento de la predicción y linealmente cuanto menor sea J_{ins} . Por otro lado, la precisión disminuye con la disminución de J_{ins} y R , especialmente en muestras de datos pequeñas, ya que los árboles que se forman están menos correlacionados dando lugar a predicciones más diferentes. Es por ello, que, al disminuir R , un aumento de J_{ins} es beneficioso ya que se necesitarán mayor cantidad de árboles para lograr la convergencia de las predicciones. En muestras de gran tamaño, se puede obtener un rendimiento óptimo con los primeros 100 árboles [51] [52].

Respecto a la relación entre variables, una disminución de R da lugar a menor capacidad de ver si varias variables están correlacionadas ya que existe una menor posibilidad de que variables relevantes sean utilizadas en un nodo para dividir los datos. Un aumento de J_{ins} mejora esta capacidad ya que aumenta la probabilidad de que en alguno de los árboles variables correlacionadas aparezcan para dividir los datos [51] [52].

El propio algoritmo de SMTS utiliza diversos valores de R y J_{ins} de una lista que introducimos para calcular el error de validación mediante un sistema que utiliza las observaciones out-of-bag (muestras que han quedado fuera del entrenamiento del algoritmo para su uso en la validación). A pesar de que este sistema tiene mayor sesgo y menor varianza que cross-validation, da lugar a resultados muy favorables en cuanto a elegir los valores de R y J_{ins} que dan lugar a un error de validación menor disminuyendo enormemente el tiempo de ejecución. En concreto, se realizan 10 combinaciones de valores de R y J_{ins} . Para obtener un resultado más robusto, se correrá el algoritmo 10 veces con los mismos valores de R y J_{ins} y así obtener el mayor número de combinaciones de estos posibles y, por lo tanto, el conjunto de valores de R y J_{ins} que dan lugar a una menor varianza en la estimación y al menor error de validación posible.

3

Resultados y análisis

A continuación, se muestran los resultados y su análisis una vez obtenidos los datos del dispositivo de seguimiento ocular a partir del protocolo descrito y tras su tratamiento y modificación acorde a lo que se requiere para cada algoritmo.

3.1 Resultados de la prueba de *eye-tracking*

Los resultados de varios de los pacientes se muestran gráficamente en el Anexo III comparando las visualizaciones de los datos y el diámetro de la pupila para cada uno de los dos textos utilizados.

A continuación, se mostrará y analizará el seguimiento ocular de uno de los pacientes a lo largo de la lectura del texto 1 (Caperucita Roja) como ejemplificación del análisis que se haría para cada uno de los pacientes.

El comportamiento de este paciente durante la lectura de los textos se puede extrapolar en otros casos, aunque, como generalidad, se ha observado una gran variabilidad entre los pacientes incluidos en el estudio. Esto se puede deber a diversas velocidades de lectura, a la opinión subjetiva de cada paciente sobre lo interesante que le resulta el texto (cuanto mayor interés mayor diámetro), la capacidad cognitiva e intelectual de base de cada paciente (cuanta menor sea mayor esfuerzo compensatorio tendrá que realizar y mayor será el diámetro pupilar), etc. [53] [54].

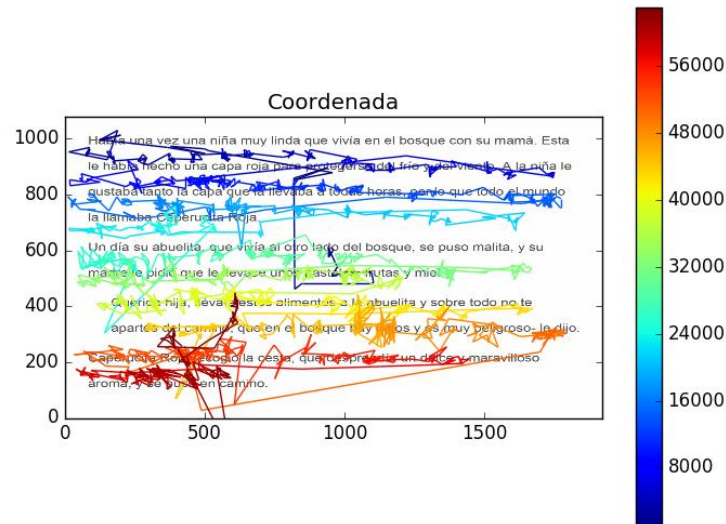


Figura 8. Ejemplo de visualización de los datos de seguimiento obtenidos para uno de los sujetos del estudio durante la lectura del cuento de Caperucita Roja a lo largo del tiempo. Estos datos corresponden a la media de las coordenadas obtenidas del ojo derecho e izquierdo.

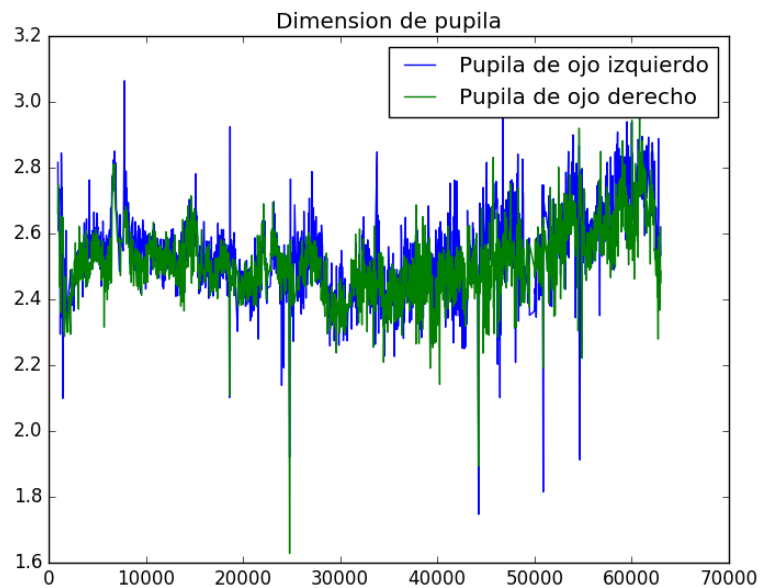


Figura 9. Visualización de los cambios en el tamaño de la pupila a lo largo del tiempo del mismo sujeto de la figura anterior durante la lectura del cuento de Caperucita Roja.

En la primera imagen podemos observar un comportamiento general de visión enfocada (lectura de las palabras y no escaneo) promovida por la tipografía relativamente pequeña del texto. El patrón se caracteriza por una continuidad en la lectura junto con fijaciones claras (nodos) en las palabras de mayor significancia (sustantivos, verbos y adjetivos). Las fijaciones se concentran en estos tipos de palabras ya que son las más complejas y el hecho de saltarse una unidad al leer es signo de facilidad de procesamiento debido a una alta predictibilidad en el contexto, por lo que se dan con palabras de mayor simplicidad.

También sobresale un inicio bastante constante en contraposición con la parte final donde aumenta el tiempo de cada fijación. Este aumento puede indicar confusión o dificultad para entender el contexto/significado. Esta dificultad sería de grado medio ya que no se observan grandes regresiones (de más de 10 caracteres). El incremento en la atención visual en la parte final también viene asociada al hecho de que el párrafo sea más corto en comparación con el resto, ya que cuanto más largo sea menor atención recibe [55] [56].

Respecto a la segunda imagen se puede observar que el tamaño de las pupilas varía entre 2 y 3 mm siendo lo normal entre 2 y 4 mm con luz y entre 4 y 8 mm en oscuridad [57], para adaptarse a las condiciones lumínicas ambientales y a la actividad que se esté realizando. Este precepto se cumple en prácticamente la totalidad de la lectura.

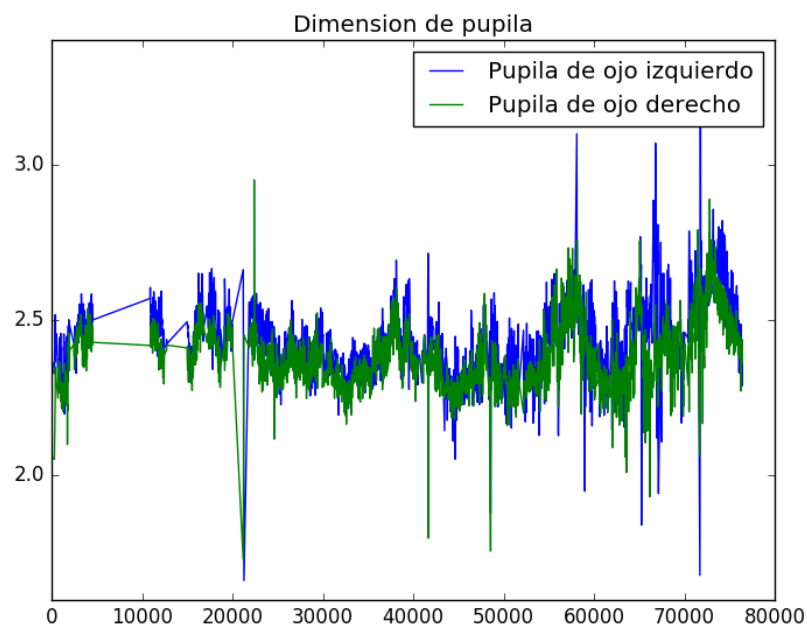


Figura 10. Visualización de los cambios en el tamaño de la pupila a lo largo del tiempo para el mismo sujeto que en las figuras anteriores durante la lectura del fragmento de Don Quijote de la Mancha.

Por otro lado, comparando la dimensión media de la pupila en la lectura del cuento de Caperucita Roja (Figura 8) y en el fragmento del Don Quijote (Figura 9), la media en el primer caso es claramente superior. Esto se debe a un comportamiento asociado a la memoria cognitiva. Según esta pauta, la predictibilidad contextual aumenta el diámetro de la pupila, especialmente en textos que nos resultan familiares como es el caso del cuento infantil [58].

En cuanto a las fluctuaciones durante la lectura de Caperucita Roja, el aumento del tamaño de la pupila se relaciona con un incremento del esfuerzo en la lectura. Así pues, observamos una tendencia al alza en el final del fragmento que, junto con el aumento en el tiempo de fijaciones que se da en esta parte también, nos indica un aumento en la concentración del paciente durante la lectura final [58].

3.2 Comparativa de los algoritmos de clasificación

Para cada uno de los algoritmos de clasificación utilizados para diferenciar las series temporales completas de pacientes con patología de aquellos sanos, se intentó conseguir la mejor combinación de parámetros que diera lugar a una precisión en la clasificación mayor del 85%.

3.2.1 SMTS

A continuación, se muestran los errores obtenidos en la clasificación con los distintos valores de R y J_{ins} utilizados durante el entrenamiento. Los errores obtenidos son de dos tipos, el global (Test Error), obtenido a partir de las muestras de prueba una vez construido el RF, y el error out-of-boostap (OOB Error), obtenido a partir de los errores individuales de cada árbol usando las muestras que no se usaron para construirlo. Estos valores se consiguieron tras la ejecución de 10 veces del algoritmo con los mismos parámetros de entrada ($R = (10,20,50,100,150,200)$ y $J_{ins} = (20,50,100,150,200,250,500,1000)$). De esta forma, además de aumentar la robustez del resultado, se pudo obtener los valores de OOB Error y Test Error para todas las combinaciones de R y J_{ins} posibles.

A partir de los datos obtenidos observamos que la combinación que da lugar a un error de test menor es para $R=200$ y $J_{ins}=100$ y 200 con una diferencia sustancial respecto al resto de valores. Con esta combinación obtenemos un error con los datos de prueba de 0,077 lo que implica una precisión del 92,30% en la clasificación correcta de los pacientes en sanos y con patología. El hecho de que los valores de ambos parámetros (R y J_{ins}) sean considerablemente altos sugiere que los conjuntos de datos obtenidos para cada paciente tienen una complejidad alta.

A continuación, calculamos los errores medios de los datos de test para cada combinación R - J_{ins} de las 10 repeticiones del entrenamiento realizadas. El resultado se muestra en las Figuras 11 y 12 donde se expone la variación del error en función del tamaño de R y J_{ins} .

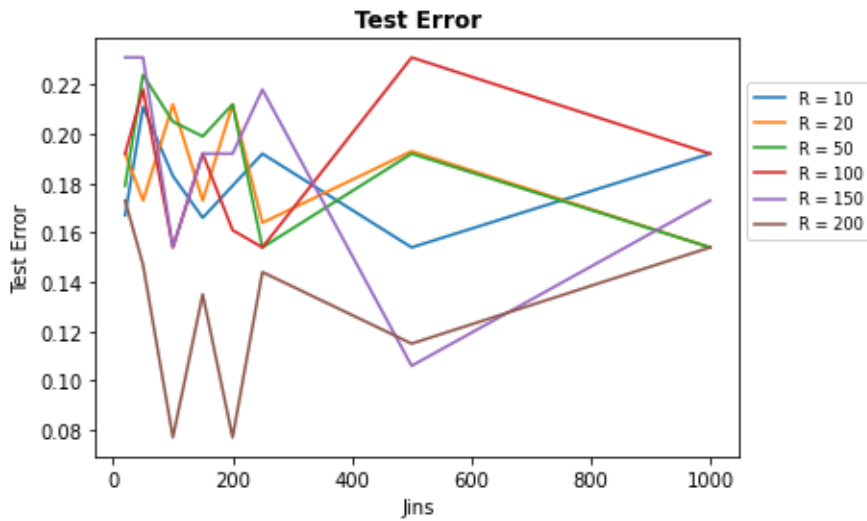


Figura 11. Representación del error obtenido con los datos de prueba con cada par de valores R-J_{ins} mostrando los valores de J_{ins} en el eje X.

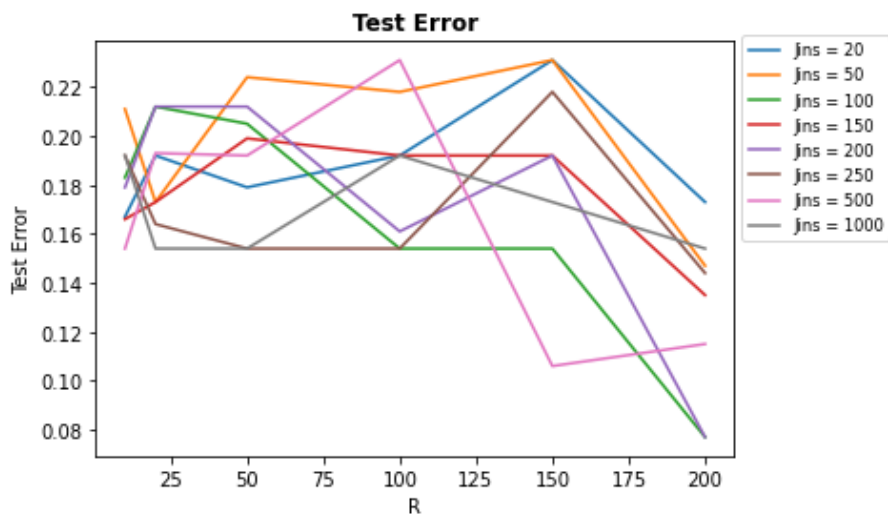


Figura 12. Representación del error obtenido con los datos de prueba con cada par de valores R-J_{ins} mostrando los valores de R en el eje X.

Se puede observar en dichas imágenes que el error se mantiene relativamente constante, aunque podemos observar claras variaciones dependientes de los valores de R y J_{ins}. El porcentaje de variación del error entre el máximo y el mínimo obtenido es de un 15,40% (error mínimo = 0.077 y error máximo = 0.231).

En cuanto a las fluctuaciones con respecto a los valores de R, se define una tendencia al alza con R < 150 y a la baja con R > 150. Esto queda claro, especialmente, en la Figura 12 donde el error disminuye drásticamente a partir de R = 150 para todos los valores de J_{ins}.

Respecto a J_{ins} , se observa un perfil sin ninguna característica distintiva lo que nos da pie a pensar que la influencia de los valores de este parámetro es mucho menor que la de los valores de R ya que el aumento y disminución del error no parece verse influenciado por la variación en los valores de J_{ins} .

3.2.1.1 Optimización de SMTS

Para intentar obtener un porcentaje de precisión en la clasificación mayor, se estudió el parámetro del retraso temporal.

En el algoritmo SMTS únicamente se estudia la relación entre un dato y el dato anterior cuando es posible que exista una relación entre el propio dato y otros descritos con mayor anterioridad. Esta posible relación de datos a intervalos mayores puede ser relevante cuando la velocidad de muestreo es alta y, por lo tanto, las diferencias entre los valores obtenidos de un momento al siguiente no son significativas.

Esta relación entre datos a distancia es lo que denominamos el retraso temporal en la codificación de entrada. En nuestro caso lo hemos estudiado eliminando primero, uno de cada dos datos y, después, uno de cada tres, de tal forma que modificamos la velocidad de muestreo de 30 Hz a 15 Hz y 7,5 Hz respectivamente.

A partir de estos datos, la precisión media obtenida usando la mejor configuración para el algoritmo SMTS conseguida en el apartado anterior ($R = 200$ y $J_{ins} = 100$ y 200) tras 10 repeticiones del entrenamiento se muestra en la tabla 5.

Tabla 5. Porcentajes de precisión en la clasificación de los datos provenientes de la residencia Domus VI con el uso del algoritmo SMTS optimizado dependiendo de la velocidad de muestreo y el número de árboles (J_{ins}) del algoritmo.

		Velocidad de muestreo	
		15 Hz	7,5 Hz
J_{ins}	100	92,90%	91%
	200	89,43%	88,21%

Así pues, podemos observar un muy ligero aumento en la precisión para el aumento de la velocidad de muestreo a 15 Hz y una disminución de un 4,43% de esta para 7,5 Hz. Estos valores sugieren que el retraso temporal en la codificación no mejora sensiblemente la precisión en la clasificación lo que puede ser debido a que las relaciones entre los datos se dan a corta distancia o a una distancia aún mayor que la estudiada.

3.2.2 LSTM

Esta red neuronal recurrente se conformó basándonos en el precepto de que cuanto más simple es la red mejor. De esta forma, buscaremos una red neuronal que nos proporcione una precisión alta sin necesidad de que sea muy compleja para, así, evitar el sobreajuste.

Por esta razón, para la estructura base de este algoritmo nos decantamos por la más simple posible: una capa de LSTM vainilla y una capa de salida con 1 neurona, ya que nuestro resultado será 1 para pacientes con patología y 0 para sanos.

Basándonos siempre en el precepto de simplicidad, los parámetros ajustables de esta red se mantuvieron todos con los valores existentes por defecto a excepción de tres de ellos: el número de unidades en la capa LSTM, el número de épocas del entrenamiento y el optimizador. De esta forma, los valores para todos los parámetros usados en la red neuronal para su optimización serán:

- *Units*: se escogieron valores relativamente bajos para evitar aumentar la complejidad de la red. Tras varias ejecuciones iniciales con distintos valores nos decantamos por los siguientes: 30, 50, 75, 100, 150, 200.
- *Epochs*: para definir los valores a utilizar se corrió el modelo con 400 épocas en un primer momento. A partir de ahí pudimos observar que a partir de 250 se producía un claro sobreajuste ya que los valores de precisión disminuían y los de pérdida de precisión aumentaban. Por esta razón, para cada valor de Units se utilizaron los siguientes números de épocas: 50, 100, 150, 200, 250.
- *Dropout_rate*: se escogió el valor por defecto, 0.0. Se decidió no evitar el sobreajuste con este parámetro ya que implica la pérdida de datos durante el entrenamiento y, en nuestro caso, no nos podíamos permitir la disminución de los datos totales ya que contamos, de inicio, con pocos datos.
- *Activation*: la función mayormente utilizada es la sigmoide. En nuestro caso, se probaron tanto la sigmoide como la Softmax escogiendo esta última tanto por la obtención de mejores resultados como con vistas al futuro de una clasificación con más de dos clases. Esta es una variación de la función sigmoidea para clasificaciones de clases múltiples que devuelve un vector con la probabilidad de pertenecer a cada clase. Esta decisión se basó en la prospección futura de utilizar este algoritmo para diferenciar las propias patologías neurodegenerativas entre ellas aumentando la cohorte de pacientes de cada una de ellas.
- *Batch_size*: para nuestro algoritmo, cada serie temporal de cada paciente la hemos dividido en cuatro para así obtener más datos para un entrenamiento más eficiente. Ya que el tiempo no es un problema para una red tan simple y con pocos datos como esta, hemos preferido primar la estabilidad del aprendizaje. Es por ello por lo que el valor de este parámetro será el tamaño de la partición más grande conseguida a partir de la división de las series temporales, 740.
- *Optimizer*: utilizamos Adam ya que es el más usado en general y, en concreto, en problemas de clasificación con datos multivariantes.
- *Loss function*: Inicialmente se comenzó el entrenamiento con categorical crossentropy que es la función utilizada en estudios como este donde tenemos

diversas clases para la clasificación. Tras obtener resultados con precisiones muy bajas, se cambió a binary crossentropy, función similar a la anterior, pero para problemas de clasificación binaria. Los resultados de precisión continuaron siendo desalentadores por lo que, finalmente, se probó con la función de mean square error con la que se obtuvieron los resultados más prometedores.

A continuación, llevamos a cabo la normalización de los datos. En este paso, observamos valores atípicos que dan lugar a errores en el entrenamiento. Es por ello por lo que estos datos fueron eliminados.

Con los datos finales, observamos que la combinación que da lugar a una pérdida de precisión menor y, por lo tanto, a una precisión mayor es para épocas = 200 y número de neuronas = 75. Con esta combinación obtuvimos una precisión del 85,71% en la clasificación correcta de los pacientes en control y estudio. Con esto podemos deducir la existencia de cierta complejidad en los propios datos y en su relación con el paciente sano o con patología ya que tanto el número de épocas necesarias para entrenar el modelo como el número de neuronas de la capa LSTM es relativamente elevado.

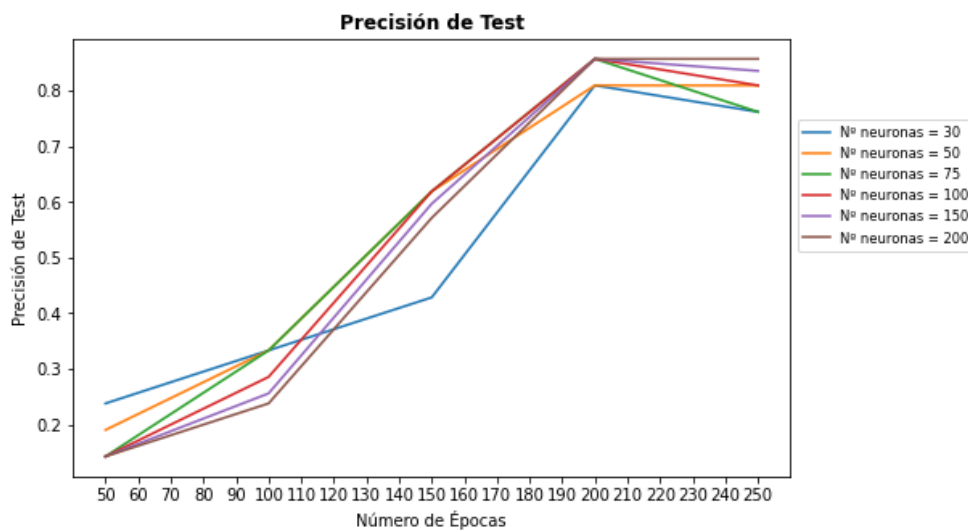


Figura 13. Representación de la precisión obtenida con los datos de prueba con cada par de valores época-nº neuronas mostrando los valores de épocas en el eje X.

En la Figura 13 podemos observar una enorme influencia del número de épocas en la precisión final del modelo. A mayor número, mayor precisión, produciéndose un salto significativo entre 100 y 200 épocas donde la precisión se llega a triplicar para ciertos valores del número de neuronas. Esta imagen también nos da una idea del momento donde comienza a estabilizarse la precisión máxima e, incluso, disminuye ligeramente. Este momento es clave para reconocer el inicio del sobreajuste y se produce cuando se alcanzan las 200 épocas. El comportamiento de este fenómeno, para la mejor configuración del algoritmo, se puede observar mejor en la Figura 14.

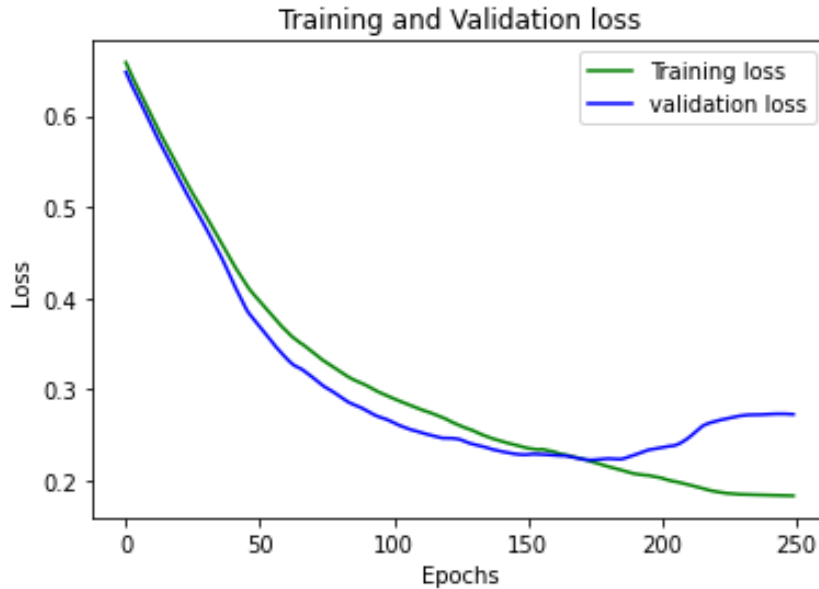


Figura 14. Pérdida de precisión durante el entrenamiento y la validación de la red neuronal recurrente LSTM formada por 75 neuronas durante 250 épocas.

A partir de las 180 épocas podemos observar un incremento en el error de validación mientras que el error de entrenamiento continúa disminuyendo. Esto es un claro indicador de sobreajuste por lo que el entrenamiento de la red se tendría que llevar a cabo hasta este valor de épocas

A diferencia del número de épocas, el de neuronas no influye prácticamente en la mejora o empeoramiento de la precisión del modelo.

3.3 Clasificación por tipo de enfermedad

En este apartado usaremos el algoritmo SMTS para tratar con los datos obtenidos del hospital de pacientes sanos, con accidentes cerebrovasculares y con traumatismos craneoencefálicos. El objetivo es alcanzar una precisión de más del 85% en la clasificación del tipo de enfermedad que sufre cada individuo.

Este estudio se realizará únicamente con el algoritmo SMTS ya que fue el que mejor resultado aportó en la etapa anterior. Así pues, la ejecución del algoritmo seguirá el mismo protocolo y los mismos parámetros descritos en la sección 3.2.1.

Mediante el uso de este algoritmo conseguimos una precisión máxima en la clasificación de individuos en cada una de las tres clases posibles de un 89,60% para los valores de $R = 50$ y $J_{ins} = 250$. Vemos que estos valores son inferiores a los requeridos para el primer conjunto de datos lo que puede ser una indicación de una menor complejidad en las relaciones de cada conjunto de datos con cada una de las clases.

A continuación, calculamos los errores medios de los datos de test para cada combinación R-J_{ins} de las 10 repeticiones del entrenamiento realizadas. El resultado se muestra en la Figura 15 donde se expone la variación del error en función de J_{ins}.

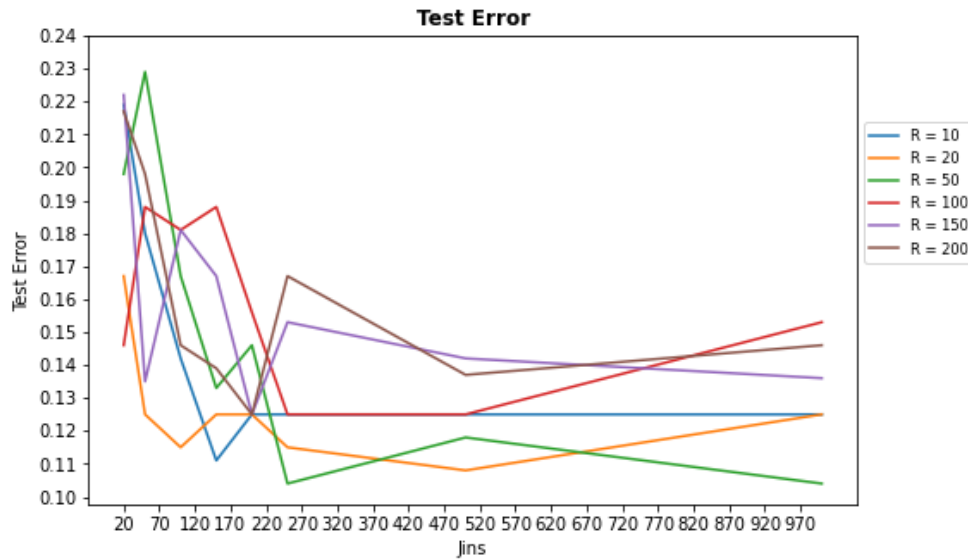


Figura 15. Representación del error obtenido con los datos de prueba procedentes del hospital con cada par de valores R-J_{ins} mostrando los valores de J_{ins} en el eje X.

El porcentaje de variación del error entre el máximo y el mínimo obtenido es de un 12,50% (error mínimo = 0.104 y error máximo = 0.229).

Podemos observar una gran influencia de los valores de J_{ins} en la clasificación correcta de los individuos. Se observa una tendencia descendente del error hasta 500 árboles. A partir de ahí el error se mantiene constante o comienza a aumentar ligeramente. Este comportamiento se da para cualquier valor de R excepto para R = 50 que alcanza el error mínimo a los 250 árboles, crece hasta los 500 y vuelve a disminuir a continuación hasta volver al error mínimo. Generalizando, valores menores de 250 en J_{ins} dan lugar a un aumento importante en el error.

Por otro lado, aunque con menor trascendencia, la gráfica nos muestra una dependencia del error respecto al valor de R. Valores bajos dan lugar a un porcentaje de precisión mayor en la clasificación que aquellos altos.

Este comportamiento contrasta con el observado en la clasificación binaria donde primaba la importancia de R frente a J_{ins}. Así pues, podemos deducir que a medida que aumentan el número de clases para clasificar a los pacientes prima la influencia del número de árboles frente a la del tamaño del alfabeto.

4

Conclusiones

El objetivo principal de este estudio ha sido desde el inicio la creación de un protocolo sencillo, rápido y cómodo para el apoyo al diagnóstico de enfermedades neurodegenerativas/neurológicas. Este protocolo lo hemos basado en el análisis de series temporales multivariantes íntegras, obtenidas a partir de la lectura de textos breves mediante un *eye-tracker*, a través de un algoritmo de aprendizaje automático.

Gracias a la colaboración de la Residencia Domus VI y del Hospital de Los Madroños, logramos la participación de 27 voluntarios para la validación de dicho protocolo.

Con la pequeña cohorte procedente de la residencia (15 voluntarios), realizamos la validación inicial del protocolo con un preprocesamiento mínimo de los datos para mantener tanto la estructura temporal de los patrones de lectura como la información derivada de ellos. Con esto, obtuvimos un porcentaje de precisión de clasificación del 92,30% para el algoritmo Multivariate Time Series Classification with Learned Discretization (SMTS), desarrollado por Mustafa Gokce Baydogan y George Runger, y del 85,71% para la red neuronal Long Short-Term Memory, implementada con Keras. Estos resultados se consiguieron con una parametrización simple pero cuyos altos valores nos sugieren que nos encontramos ante unos datos que se interrelacionan de forma compleja.

Nos decantamos por la red SMTS para la siguiente fase ya que, pudiendo optimizar la red neuronal para obtener un resultado de precisión mayor este no es el objetivo de este estudio si no que el objetivo principal es la demostración de la viabilidad del sistema de clasificación a partir de series temporales de lectura con el sistema más simple posible. En cuanto a la clasificación de los pacientes por tipo de enfermedad neuronal, se obtuvo una precisión máxima del 89,60% para las tres clases estudiadas (control, accidente cerebrovascular y traumatismo craneoencefálico).

A pesar de estos buenos resultados, debemos tener en cuenta inconvenientes como son: el elevado coste del dispositivo de seguimiento ocular, lo que da lugar a que la tecnología asociada no este, todavía, muy extendida; y el hecho de que el estudio se haya realizado con una muestra pequeña lo que nos impide extrapolar el resultado al 100% a la población

general. Además, hay que tener en cuenta que, en caso de usar este protocolo en el mundo real, nos encontraremos con pacientes con características que interfieran en la clasificación correcta de su dolencia como pueden ser el nivel educativo, el grado de afectación neuronal de la patología, patologías secundarias de afectación visual o neuronal, capacidad lectora, etc.

Con todo, este prometedor estudio sienta las bases para la implementación de este sistema de recomendación de diagnóstico para patologías neurodegenerativas, e incluso patologías de afectación neurológica general, con un porcentaje de acierto de un 92,30% para el primer conjunto de enfermedades y de un 89,60% para el segundo usando un protocolo muy simple y el algoritmo de aprendizaje automático SMTS.

5

Trabajo futuro

Este estudio se presenta como una investigación inicial del uso del seguimiento ocular como sistema para la recomendación de diagnósticos en enfermedades neurodegenerativas. Su implementación real en el ámbito profesional pasa por un trabajo más exhaustivo en el que se incluye:

- La obtención de una mayor cantidad de datos: como mínimo 100 pacientes distintos con características propias diversas, como, por ejemplo, patologías secundarias variadas (TDH, Síndrome de Down, traumatismo craneoencefálico, daltonismo, cataratas, etc.).
- La parametrización más exhaustiva del algoritmo seleccionado para ser incluido en el protocolo. El estudio de la variación del `batch_size`, la velocidad de muestreo, la exploración más exhaustiva del retraso temporal, las distintas alternativas de optimización y codificación, etc. y su influencia en la precisión final.
- El estudio de la contribución diferenciada de la posición de la mirada y el tamaño de la pupila en la clasificación.

Por otro lado, en este trabajo no se pudo abordar por falta de tiempo y por la situación sanitaria actual la posibilidad de volver más específico el algoritmo de clasificación para, no solo identificar si el paciente sufre alguna enfermedad neurodegenerativa, si no en qué grado está. Esto junto con la introducción de otras enfermedades de afectación neuronal como son TDAH, trastornos del espectro autista y esclerosis múltiple, serían el siguiente paso para completar el sistema.

Además, para conseguir una implementación total, habría que explorar otros dispositivos más económicos, así como crear una aplicación con interfaz gráfica *user-friendly* que integrara la obtención de los datos del *eye-tracker* en tiempo real y el algoritmo de clasificación para así, a partir de la lectura del texto, obtener directamente la recomendación diagnóstica en pantalla. De esta forma, la transferencia industrial sería posible de forma más completa ya que se presentaría como un protocolo sencillo, rápido, no intrusivo y económico. Así pues, como paso final se tendría que crear la aplicación integrando todas las tecnologías aquí estudiadas y validadas y repetir este mismo estudio, junto con las mejoras

mencionadas arriba, con los datos obtenidos de dispositivos más económicos como el Smart Eye AI-X o el Gazepoint GP3.

Finalmente, como elemento de avance para un estudio más a largo plazo, se podría introducir esta idea y sistema en una aplicación móvil usando la cámara del dispositivo (smartphone, Tablet u ordenador) para obtener los datos de lectura en tiempo real. Esta posibilidad sería interesante para la identificación temprana de patologías en niños, realizar un seguimiento 24/7 en pacientes con diagnóstico de tal forma que se pudiera corregir la medicación lo más rápido posible en caso de desarrollo de la enfermedad y, también, en personas de avanzada edad integrado como un sistema de autocuidado de tal forma que, en caso de desarrollarse cualquier enfermedad neurológica se pudiera frenar/tratar lo más rápido posible gracias a un diagnóstico temprano.

Bibliografía

- [1] T. C. Benjamin y G. L. Steven, «Best practices in eye tracking research,» *International Journal of Psychophysiology*, vol. 155, p. 49–62, 2020.
- [2] K. Harezlak y P. Kasprowski, «Application of eye tracking in medicine: A survey, research issues and challenges,» *Computerized Medical Imaging and Graphics*, vol. 65, pp. 176-190, 2018.
- [3] O. Ferhat y F. Vilariño, «Low Cost Eye Tracking: The Current Panorama,» *Computational Intelligence and Neuroscience*, vol. 3, nº 2, pp. 1-14, 2016.
- [4] H. Singh y J. Singh, «Human eye tracking and related issues: A review,» *International Journal of Scientific and Research Publications*, vol. 2, nº 9, pp. 1-9, 2012.
- [5] M. Wedel y R. Pieters, «A review of eye-tracking research in marketing,» *Review of marketing research*, pp. 123-147, 2017.
- [6] A. J. Larrazabal, C. G. Cena y C. E. Martínez, «Video-oculography eye tracking towards clinical applications: A review,» *Computers in biology and medicine*, vol. 108, pp. 57-66, 2019.
- [7] J. Z. Lim, J. Mountstephens y J. Teo, «Emotion recognition using eye-tracking: taxonomy, review and current challenges,» *Sensors*, vol. 20, nº 8, p. 2384, 2020.
- [8] V. Skaramagkas, G. Giannakakis, E. Ktistakis, D. Manousos, I. Karatzanis, N. Tachos y M. Tsiknakis, «Review of eye tracking metrics involved in emotional and cognitive processes,» *IEEE Reviews in Biomedical Engineering*, 2021.
- [9] A. Wolf, K. Ueda y Y. Hirano, «Recent updates of eye movement abnormalities in patients with schizophrenia: A scoping review,» *Psychiatry and clinical neurosciences*, vol. 75, nº 3, pp. 82-100, 2021.
- [10] M. Płużyczka, «The First Hundred Years: a History of Eye Tracking as a Research Method,» *Applied Linguistics Papers*, p. 101–116, 25 Abril 2018.
- [12] H. Chennamma y X. Yuan, «A survey on eye-gaze tracking,» *Indian Journal of Computer Science and Engineering (IJCSSE)*, pp. Vol. 4, No.5, Oct-Nov, 2013.
- [13] K. Wang, R. Zhao, H. Su y Q. Ji, «Generalizing Eye Tracking with Bayesian Adversarial Learning,» *2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 11899-11908, 2019.
- [14] M. G. Lorenzo, «Integración de tecnologías de adquisición y procesamiento de señal ocular para su uso en aplicaciones biomédicas,» *Trabajo Fin de Grado*, 2018.
- [15] A. Tsukada, M. Shino, M. Devyver y T. Kanade, «Illumination-free gaze estimation method for first-person vision wearable device,» de *IEEE International Conference on Computer Vision Workshops*, Barcelona, 2011.

- [16] X. Zhang, Y. Sugano, M. Fritz y A. Bulling, «Appearance-based gaze estimation in the wild,» *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 4511-4520, 2015.
- [17] L. Tao, Q. Wang, D. Liu, J. Wang, Z. Zhu y L. Feng, «Eye tracking metrics to screen and assess cognitive impairment in patients with neurological disorders,» *Neurological Sciences*, nº 41, p. 1697–1704, 2020.
- [18] «World Health Organization,» 2020. [En línea]. Available: <https://www.who.int/news-room/fact-sheets/detail/the-top-10-causes-of-death#:~:text=The%20top%20global%20causes%20of,birth%20asphyxia%20and%20birth%20trauma%2C>.
- [19] Consejo Interterritorial del Sistema Nacional de Salud, «Estrategia en Enfermedades Neurodegenerativas del Sistema Nacional de Salud,» Ministerio de Sanidad, Servicios Sociales e Igualdad, Madrid, 2016.
- [20] A. E. d. E. Neurodegenerativas, «Estudio sobre las enfermedades neurodegenerativas en España y su impacto económico y social,» *Universidad Complutense de Madrid*, 2016.
- [21] «NeuroWikia,» 13 04 2021. [En línea]. Available: <http://www.neurowikia.es/content/anatomia-patologica-de-la-enfermedad-de-alzheimer>.
- [22] A. Guimerà, X. Gironès y F. Cruz-Sánchez, «Actualización sobre la patología de la enfermedad de Alzheimer,» *Revista Española de Patología*, vol. 35, pp. 21-48, 2002.
- [23] D. Furcila, J. DeFelipe y L. Alonso-Nanclares, «A study of amyloid- β and phosphotau in plaques and neurons in the hippocampus of Alzheimer's disease patients,» *Journal of Alzheimer's Disease*, vol. 64, pp. 417-435, 2018.
- [24] M. T. Gerrero Díaz, M. C. Macías Montero, E. F. Prado, P. A. Muñoz, M. V. Hernández Jiménez y L. J. Duarte García, *Tratado de geriatría para residentes*, Madrid: International Marketing & Communication, 2007.
- [25] R. Martínez-Fernández, C. Gasca-Salas, A. Sánchez-Ferro y J. A. Obeso, «Actualización en la enfermedad de Parkinson,» *Revista Médica Clínica Las Condes*, vol. 27, pp. 363-379, 2016.
- [26] «Clínica Universidad de Navarra,» 13 04 2021. [En línea]. Available: <https://www.cun.es/enfermedades-tratamientos/enfermedades/demencia-frontotemporal>.
- [27] Á. M. Iragorri Cucalón, «Demencia frontotemporal,» *Revista colombiana de psiquiatría*, pp. 139-156, 2007.
- [28] P. Lillo y C. Leyton, «Demencia frontotemporal, cómo ha resurgido su diagnóstico,» *Revista Clínica Médica Las Condes*, vol. 27, pp. 309-318, 2016.
- [29] G. Faggioni Sánchez, G. Herbozo Alvarado y e. al., «Demencia por cuerpos de Lewy, un reto diagnóstico,» *Revista Ecuatoriana de Neurología*, vol. 27, pp. 2631-2581, 2018.
- [30] J. Toro, «Demencia con cuerpos de Lewy,» *Acta Neurol Colomb*, vol. 26, pp. 78-80, 2010.

- [31] U. P. Mosimann, R. M. Müri y e. al., «Saccadic eye movement changes in Parkinson's disease dementia and dementia with Lewy bodies» *Brain*, vol. 128, pp. 1267-1276, 2005.
- [32] C. McDonald Connor y e. col., «Individual differences in fifth graders' literacy and academic language predict comprehension monitoring development: An eye movement study,» *Scientific Studies of Reading*, vol. 19, nº 2, pp. 114-134, 2015.
- [33] A. Godroid y e. col., «Incidental vocabulary learning in a natural reading context: an eye-tracking study,» *Bilingualism: Language and Cognition*, vol. 21, nº 3, p. 563-584, 2018.
- [34] S. Garbutt, M. Alisa y e. al., «Pculomotor function in frontotemporal lobar degeneration, related disorders and Alzheimer's disease,» *Brain*, vol. 131, pp. 1268-1281, 2008.
- [35] A. G. Douglass, «Eye movements in neurocognitive disorders and frontotemporal dementia,» *University of Melbourn*, 2016.
- [36] P.-H. Tseng, I. G. M. Cameron, G. Pari, J. N. Reynolds, D. P. Muñoz y L. Itti, «High-throughput classification of clinical populations from natural viewing eye movements,» *Journal of Neurology*, vol. 260, pp. 275-284, 2013.
- [37] Tobii Technology, «User Manual,» 2013.
- [38] A. M. Mento, «Different Kinds of Eye Tracking Devices,» 2020.
- [39] R. Zemblys, D. C. Niehorster, O. Komogortsev y K. Holmqvist, «Using machine learning to detect events in eye-tracking data,» *Behavior Research Methods*, vol. 50, nº 1, pp. 160-181, 2018.
- [40] D. Rozado, J. S. Agustin, F. B. Rodriguez y P. Varona, «Gliding and saccadic gaze gesture recognition in real time,» *ACM Transactions on Interactive Intelligent Systems*, vol. 1, nº 2, pp. 1-27, 2012.
- [41] D. Rozado, F. B. Rodriguez y P. Varona, «Optimizing Hierarchical Temporal Memory for Multivariable Time Series,» *Lect. Notes Comput. Sci.*, vol. 6353, pp. 506-518, 2010.
- [42] X. Zhou, X. Gao, J. Wang, H. Yu, Z. Wang y Z. Chi, «Eye tracking data guided feature selection for image classification,» *Pattern Recognition*, vol. 63, pp. 56-70, 2017.
- [43] D. Rozado, F. B. Rodriguez y P. Varona, «Low cost remote gaze gesture recognition in real time,» *Applied Soft Computing*, vol. 12, nº 8, pp. 2072-2084, 2012.
- [44] M. Venkatachalam, «Recurrent Neural Networks,» 1 Marzo 2019. [En línea]. Available: <https://towardsdatascience.com/recurrent-neural-networks-d4642c9bc7ce>. [Último acceso: 1 Diciembre 2020].
- [45] IBM Cloud Education, «Recurrent Neural Network,» IBM, 14 Septiembre 2020. [En línea]. [Último acceso: 1 Diciembre 2020].
- [46] R. D. Hager y D. Gregory, «Deep learning: RNNs and LSTM,» de *Handbook of Medical Image Computing and Computer Assisted Intervention*, The Elsevier and MICCAI Society Book Series, 2020, pp. 503-519.
- [47] Schmidhuber, S. Hochreiter y Jürgen, «Long Short-Term Memory,» *Neural Computation*, vol. 9, nº 8, pp. 1735-1780, 1997.
- [48] J. Brownlee, «How to Tune LSTM Hyperparameters with Keras for Time Series Forecasting,» *Machine Learning Mastery*, 28 Agosto 2020. [En línea]. Available:

<https://machinelearningmastery.com/tune-lstm-hyperparameters-keras-time-series-forecasting/>. [Último acceso: 18 Febrero 2021].

- [49] J. Brownlee, «How to Reduce Overfitting With Dropout Regularization in Keras,» *Machine Learning Mastery*, 25 Agosto 2020. [En línea]. Available: <https://machinelearningmastery.com/how-to-reduce-overfitting-with-dropout-regularization-in-keras/>. [Último acceso: 18 Febrero 2021].
- [50] *Keras API reference*.
- [51] P. Probst, M. Wright y A.-L. Boulesteix, «Hyperparameters and Tuning Strategies for Random Forest,» *arXiv*, 2019.
- [52] E. Scornet, «TUNING PARAMETERS IN RANDOM FORESTS,» *ESAIM: PROCEEDINGS AND SURVEYS*, vol. 60, pp. 144-162, 2018.
- [53] O. Barral, I. Kosunen y G. Jacucci, «Influence of reading speed on pupil size as a measure of perceived relevance,» Helsinki, 2014.
- [54] E. L. Granholm, M. S. Panizzon y e. al., «Pupillary Responses as a Biomarker of Early Risk for Alzheimer's Disease,» *Journal of Alzheimer's Disease*, vol. 56, nº 4, pp. 1419-1428, 2017.
- [55] Tobii, «Investigación de Lectura y Lenguaje,» [En línea]. Available: <https://www.tobiiipro.com/es/aplicaciones/investigacion-cientifica/investigacion-lectura-lenguaje/>. [Último acceso: 15 Mayo 2021].
- [56] F. Martínez Gutiérrez, «Eyetracking y usabilidad: claves de investigación en los procesos de lectura en línea,» *Revista de Comunicación de la SEECI*, nº 16, pp. 98-118, 2008.
- [57] R. H. Spector, «Chapter 58: The pupils,» de *Clinical Methods: The History, Physical, and Laboratory Examinations*, Boston, Butterworths, 1990.
- [58] G. Fernández, J. Biondi, S. Castro y O. E. Agamennoni, «Pupil size behavior during on line processing of sentences,» *Journal of Integrative Neuroscience*, vol. 15, nº 4, pp. 485-496, 2016.
- [59] R. A. Armstrong, «Visual signs and symptoms of dementia with Lewy bodies,» *Clinical and Experimental Optometry*, vol. 95, pp. 621-630, 2012.
- [60] D. Fortunato, J. Sierra, I. J. Sierra, C. A. Caicedo Montaña, J. A. Mora Salazar y C. Tramontini Jens, «Anatomía Básica de los Ganglios Basales,» *Revista médica sanitas*, vol. 22, nº 2, pp. 66-71, 2019.
- [61] J. Ure y H. D'Onofrio, «Los movimientos oculares en la práctica neuropsiquiátrica,» *Revista Argentina de Clínica Neuropsiquiátrica*, vol. 7, nº 4, 1999.
- [62] J. Court, «Locus Coeruleus: aspectos fisiológicos y clínicos,» *Universidad Católica de Chile*, vol. 23, 1997.
- [63] M. Argilés Sans y M. Garabito Montoya, «Evaluación y tratamiento de los movimientos sacádicos oculares,» 2016.
- [64] C. Gutiérrez-de Piñeres Botero, «Neurobiología de los movimientos oculares,» de *Registro de movimientos oculares con el eye tracker Mobile eye XG*, Bogotá, Editorial Universidad Católica de Colombia, 2019, pp. 31-34.
- [65] V. Díaz Herráez, *Estudio de la precisión de los dispositivos de eye-tracking para la*, Universidad Autónoma de Madrid, 2017.

ANEXO I

Textos utilizados en el estudio

Había una vez una niña muy linda que vivía en el bosque con su mamá, que le había hecho una capa roja para protegerse del frío y el viento. A la niña le gustaba tanto la capa que la llevaba a todas horas, por lo que todo el mundo la llamaba Caperucita Roja.

Un día, su abuelita que vivía al otro lado del bosque se puso malita y su madre le pidió que le llevase unos pasteles, frutas y miel.

- Querida hijita, llévale estos alimentos a la abuelita y sobre todo no te apartes del camino, ya que en el bosque hay lobos y es muy peligroso – le dijo Caperucita recogió la cesta con los pasteles, la fruta y la miel y se puso en camino.

Figura 16. Imagen del primer texto utilizado en el experimento para obtener los datos de seguimiento ocular.

No comas ajos ni cebollas, porque no saquen por el olor tu villanería. Anda despacio; habla con reposo; pero no de manera que parezca que te escuchas a ti mismo, que toda afectación es mala. Come poco y cena más poco; que la salud de todo el cuerpo se fragua en la oficina del estómago. Se templado en el beber, considerando que el vino demasiado ni guarda secreto, ni cumple palabra. Ten en cuenta, Sancho, de no mascar a dos carrillos, ni de eructar delante de nadie.

Entre los pecados mayores que los hombres comenten, Sancho, aunque algunos dicen que es la soberbia, yo digo que es el desagradecimiento. Cada uno es como Dios le hizo y aún peor muchas veces.

Figura 17. Imagen del segundo texto utilizado en el experimento para obtener los datos de seguimiento ocular.

ANEXO II

Protocolo de obtención de datos de seguimiento ocular

Material:

- Ordenador con la aplicación instalada.
- Pantalla de 23,8 pulgadas.
- Dispositivo de seguimiento ocular Tobii.
- Micrófono.
- Cables conectores varios (HDMI y de alimentación).
- Transportador de ángulos.
- Luxómetro.
- Soporte para el dispositivo.
- Mesa.
- Silla con altura regulable.

Descripción:

Previamente al inicio de la prueba procederemos al montaje y encendido del ordenador.

El dispositivo y su posición final se muestran en las siguientes imágenes:

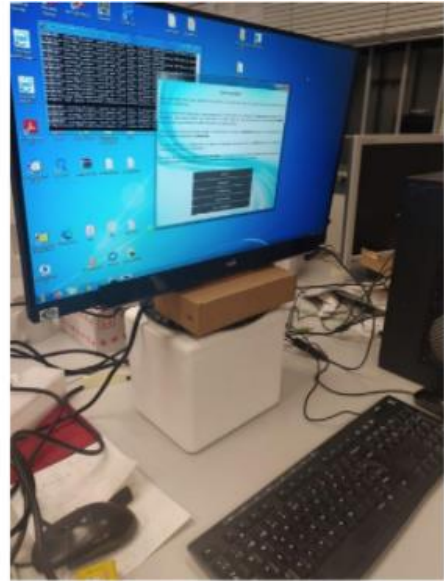


Figura 18. Disposición espacial del sistema de medición de seguimiento ocular para la realización del experimento.

Una vez iniciado se seguirán los siguientes pasos para iniciar la sesión:

- Se introducirá la siguiente contraseña: 1gnb3r0\$ (es un cero, no una O).
- Se abrirá la carpeta de TFG-Maria_Godin que se encuentra en el escritorio.
- Se abre la carpeta llamada WinPython-32bit-2.7.10.3.
- Se ejecuta el archivo WinPython Command Prompt clicando en ejecutar en caso de saltar una ventana emergente.

- En la ventana negra que aparecerá se copiarán los siguientes comandos en este orden y pulsando la tecla enter tras cada uno de ellos.
 1. `cd ..\..\Aplicacion\TFG-code`
 2. `python application.py`

De esta forma iniciaremos la aplicación para comenzar con el procedimiento.

A continuación, se describirán los pasos una vez el paciente entre en la consulta:

1. Se situará al paciente cómodamente en la silla siguiendo el siguiente esquema:

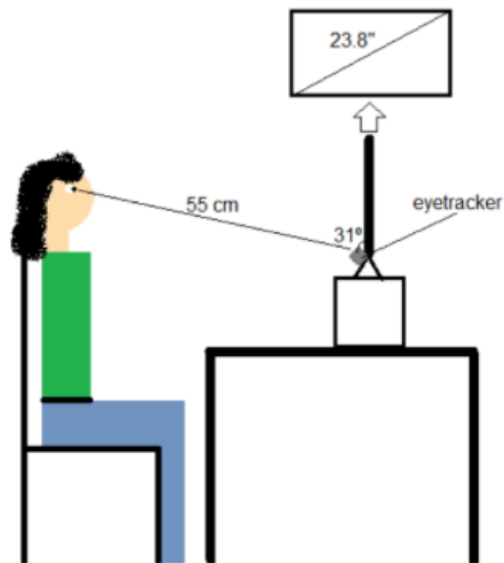


Figura 19. Esquema de la posición del voluntario respecto al dispositivo de seguimiento ocular a la hora de comenzar el experimento.

De esta forma los ojos del paciente quedarán a la misma altura que el centro de la pantalla que utilizaremos y a una distancia de 55cm del dispositivo de seguimiento ocular situado a 31o respecto a la pantalla.

2. Se le explicará con detenimiento el ejercicio hasta que consideremos la total comprensión por su parte.
3. Se abrirá el fichero2.png de la carpeta “Reading files” situada en la subcarpeta “TFG-code” de la carpeta “Aplicacion”.
4. Se posicionará el luxómetro situando la parte medidora a la altura del ojo derecho y, posteriormente, del ojo izquierdo. Los valores de los lúmenes se apuntarán en el Excel titulado “lúmenes y edad” al igual que la edad del paciente. Dicho fichero se encuentra en la carpeta de nombre “TFM_marta” que se encuentra en el escritorio.

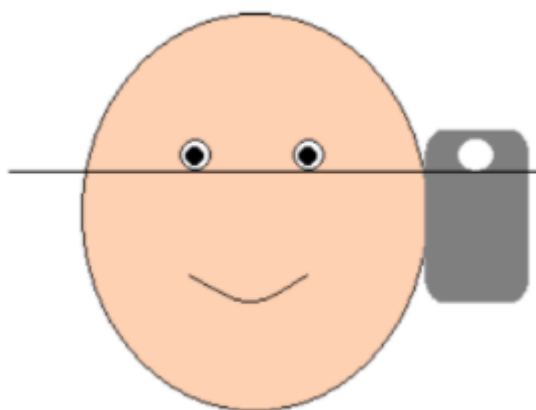


Figura 20. Esquema del posicionamiento del luxómetro a la hora de la medición de la luz ambiental previa al experimento.

5. Se iniciará la aplicación.
 - a. Usuario:
 - i. Como nombre de usuario se pondrá: paciente
 - ii. Como apellido de usuario se pondrá el número de paciente correspondiente comenzando en 1 hasta 100.
 - iii. En la edad se pondrá la edad correspondiente al paciente.
 - b. Configuración:
 - i. Se escogerá el dispositivo Tobii como EyeTracker.
 - ii. Se escogerá el único micrófono disponible.
 - iii. Se harán dos pruebas por paciente con dos textos distintos. En la casilla de "Fichero" se escogerá en primer lugar el fichero1.png al final de la dirección que se encuentra en dicha casilla, y en segundo lugar el fichero7.png.

6. Se iniciará la prueba pulsando el botón de "Realizar prueba":
 - a. Calibración: Nos aseguraremos de que durante la fase de calibración y la realización general de la prueba el paciente mueve únicamente los ojos manteniendo la cabeza estática.

Para iniciar la calibración tenemos que observar en la pantalla previa que los dos puntos están verdes y que no cambien a color rojo durante períodos prolongados o parpadeen durante un tiempo prolongado.

Se considerará válida la calibración siempre que la dispersión en la precisión y exactitud de los puntos obtenidos no sea muy grande. Un ejemplo de ello es el siguiente:



Figura 21. Ejemplo de calibración.

En caso de que no aparezcan algunos de los puntos o que la dispersión de los puntos sea muy alta y la precisión muy baja se repetirá la calibración pulsando la tecla R hasta obtener un resultado adecuado. En caso de que la calibración sea correcta se pulsará la barra espaciadora para pasar a la validación.

- b. Validación: se realizará igual que la calibración. Se siguen las instrucciones que aparecerán en la pantalla. Al terminarla se presionará la barra espaciadora para iniciar la prueba.
- c. Prueba: se darán 120 segundos para leer el texto. Una vez finalizada se cerrará la ventana y se volverá de forma automática al menú de inicio.
- d. Pulsar en obtener resultados en el menú de inicio y cerrar las ventanas emergentes. En estas ventanas se muestra gráficamente los resultados obtenidos.

7. Se realizará otra vez el proceso desde el punto 5.iii eligiendo esta vez el fichero7.png.

ANEXO III

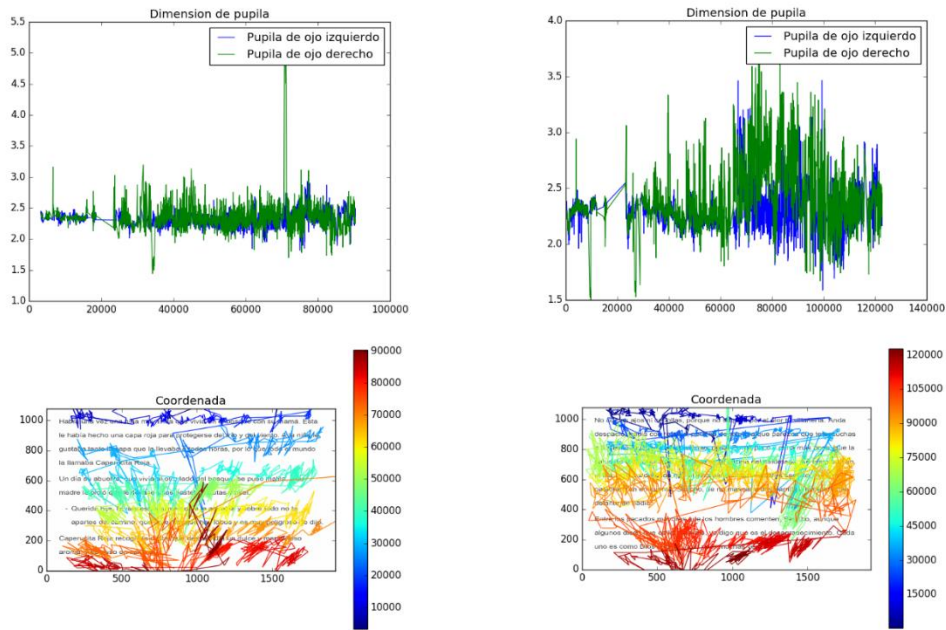


Figura 22. Comparativa de los resultados obtenidos a partir de los datos de seguimiento ocular en la prueba de lectura de dos textos del paciente B_P. En el lado izquierdo, tenemos la dimensión de la pupila (arriba) y el gráfico de visualización del seguimiento ocular (abajo) para el texto de Caperucita Roja y, en el lado derecho, las mismas imágenes para el fragmento de Don Quijote.

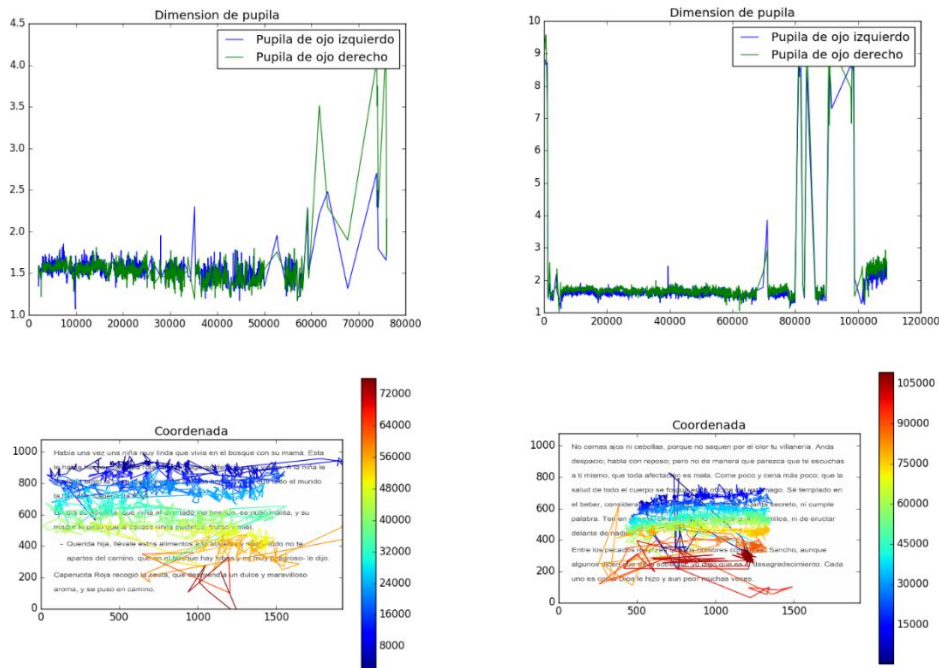


Figura 23. Comparativa de los resultados obtenidos a partir de los datos de seguimiento ocular en la prueba de lectura de dos textos del paciente C_P. En el lado izquierdo, tenemos la dimensión de la pupila (arriba) y el gráfico de visualización del seguimiento ocular (abajo) para el texto de Caperucita Roja y, en el lado derecho, las mismas imágenes para el fragmento de Don Quijote.

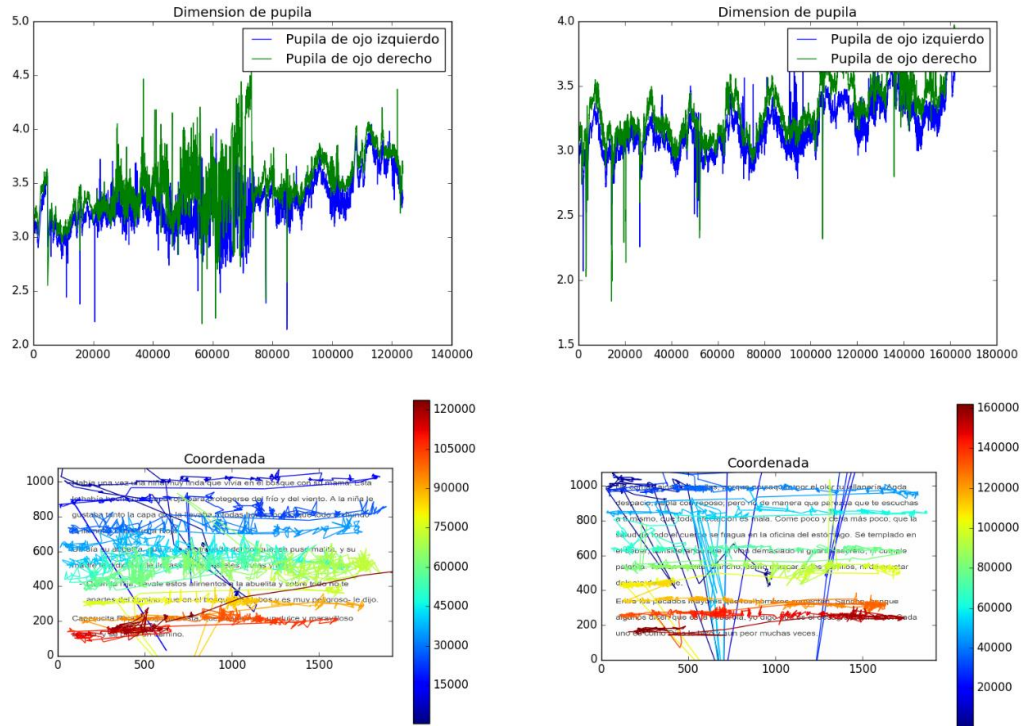


Figura 24. Comparativa de los resultados obtenidos a partir de los datos de seguimiento ocular en la prueba de lectura de dos textos del paciente D.L. En el lado izquierdo, tenemos la dimensión de la pupila (arriba) y el gráfico de visualización del seguimiento ocular (abajo) para el texto de Caperucita Roja y, en el lado derecho, las mismas imágenes para el fragmento de Don Quijote.

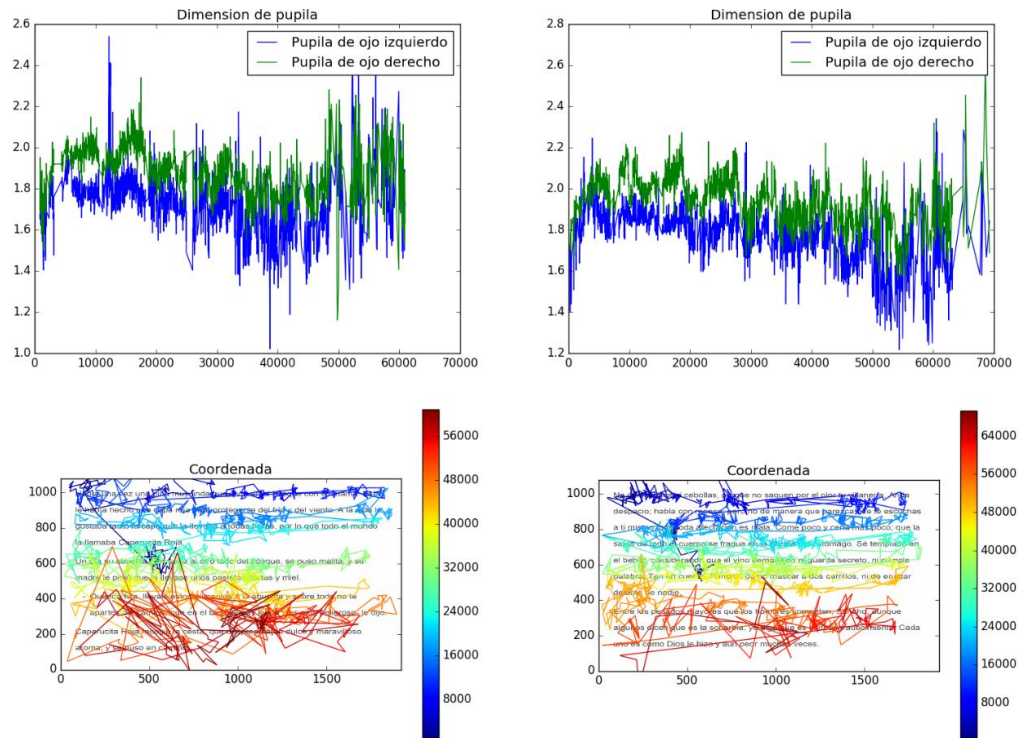


Figura 25. Comparativa de los resultados obtenidos a partir de los datos de seguimiento ocular en la prueba de lectura de dos textos del paciente F.L. En el lado izquierdo, tenemos la dimensión de la pupila (arriba) y el gráfico de visualización del seguimiento ocular (abajo) para el texto de Caperucita Roja y, en el lado derecho, las mismas imágenes para el fragmento de Don Quijote.

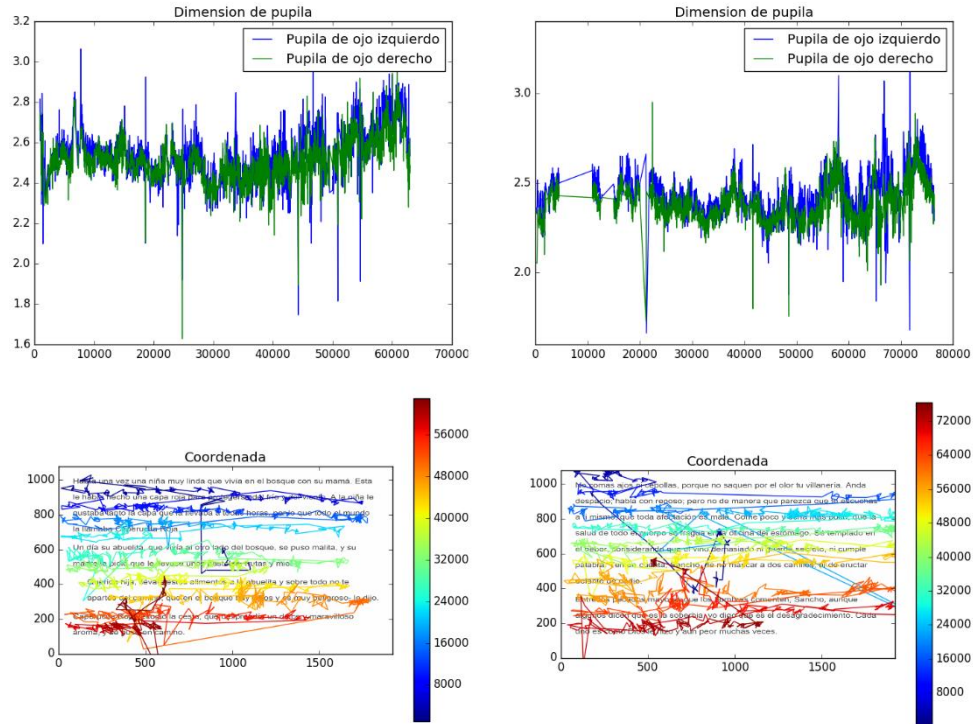


Figura 26. Comparativa de los resultados obtenidos a partir de los datos de seguimiento ocular en la prueba de lectura de dos textos del paciente G_D. En el lado izquierdo, tenemos la dimensión de la pupila (arriba) y el gráfico de visualización del seguimiento ocular (abajo) para el texto de Caperucita Roja y, en el lado derecho, las mismas imágenes para el fragmento de Don Quijote.

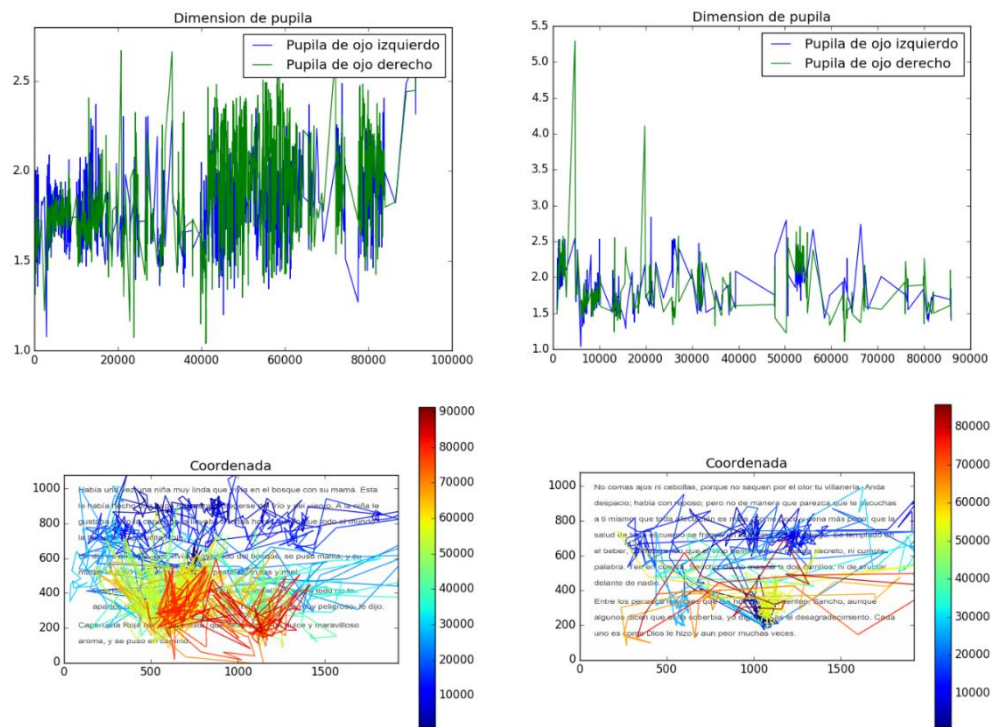


Figura 27. Comparativa de los resultados obtenidos a partir de los datos de seguimiento ocular en la prueba de lectura de dos textos del paciente G_J. En el lado izquierdo, tenemos la dimensión de la pupila (arriba) y el gráfico de visualización del seguimiento ocular (abajo) para el texto de Caperucita Roja y, en el lado derecho, las mismas imágenes para el fragmento de Don Quijote.

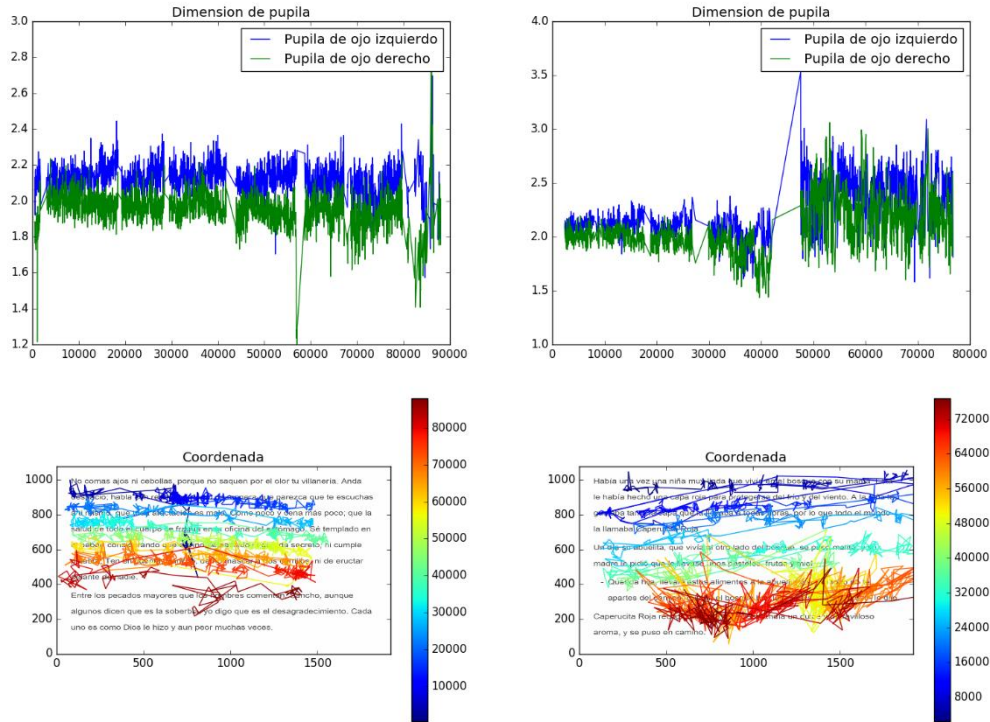


Figura 28. Comparativa de los resultados obtenidos a partir de los datos de seguimiento ocular en la prueba de lectura de dos textos del paciente G_N. En el lado izquierdo, tenemos la dimensión de la pupila (arriba) y el gráfico de visualización del seguimiento ocular (abajo) para el texto de Caperucita Roja y, en el lado derecho, las mismas imágenes para el fragmento de Don Quijote.

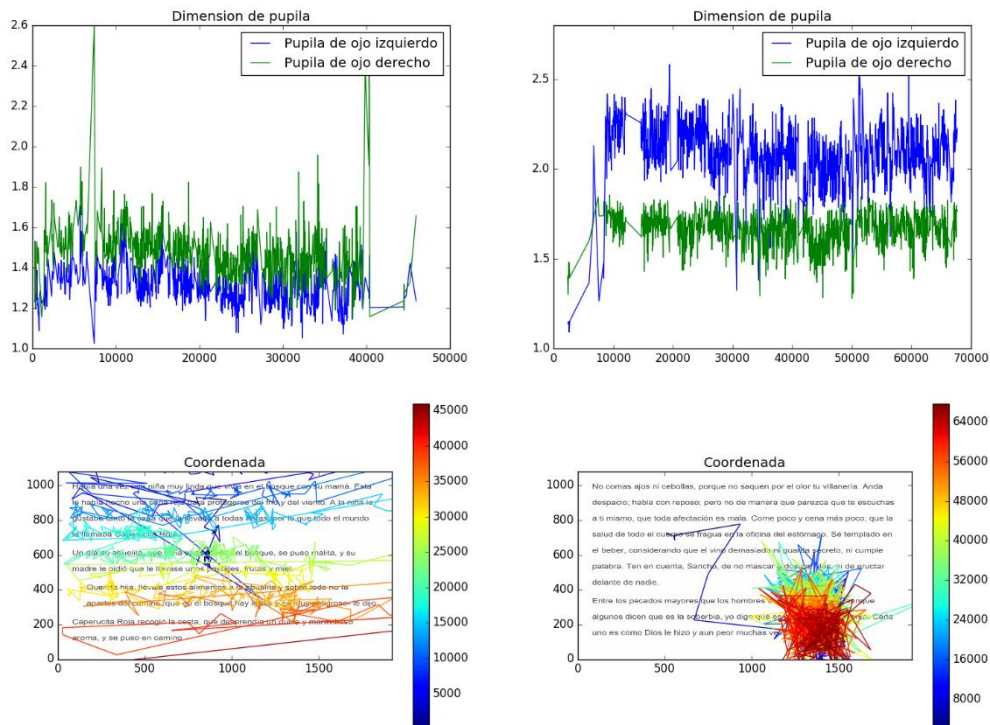


Figura 29. Comparativa de los resultados obtenidos a partir de los datos de seguimiento ocular en la prueba de lectura de dos textos del paciente G_HJ. En el lado izquierdo, tenemos la dimensión de la pupila (arriba) y el gráfico de visualización del seguimiento ocular (abajo) para el texto de Caperucita Roja y, en el lado derecho, las mismas imágenes para el fragmento de Don Quijote.

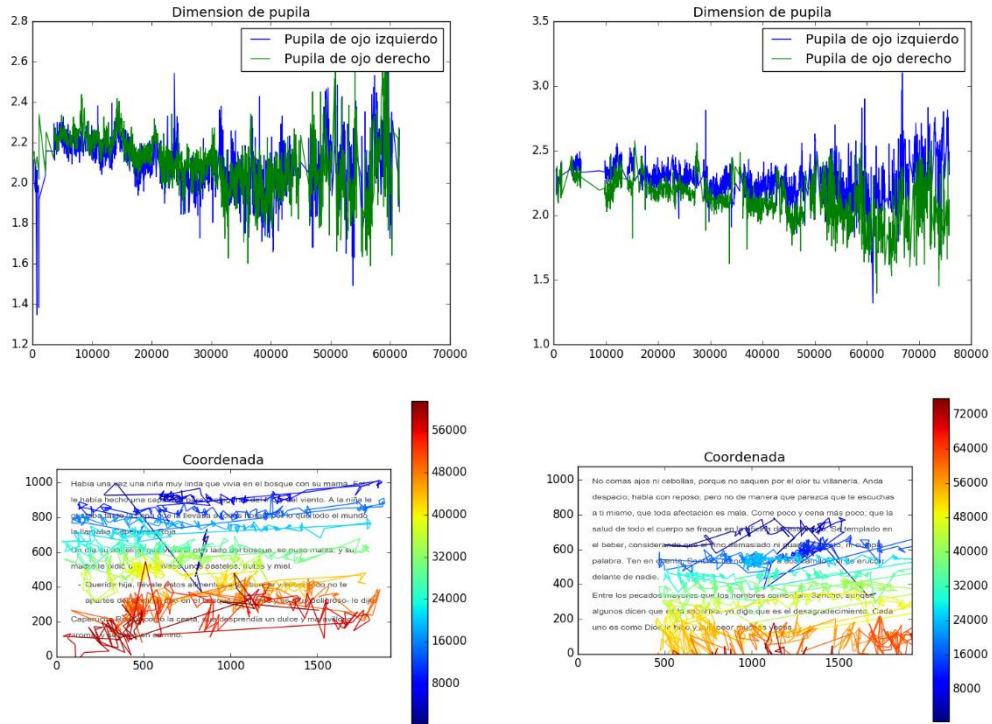


Figura 30. Comparativa de los resultados obtenidos a partir de los datos de seguimiento ocular en la prueba de lectura de dos textos del paciente M_E. En el lado izquierdo, tenemos la dimensión de la pupila (arriba) y el gráfico de visualización del seguimiento ocular (abajo) para el texto de Caperucita Roja y, en el lado derecho, las mismas imágenes para el fragmento de Don Quijote.

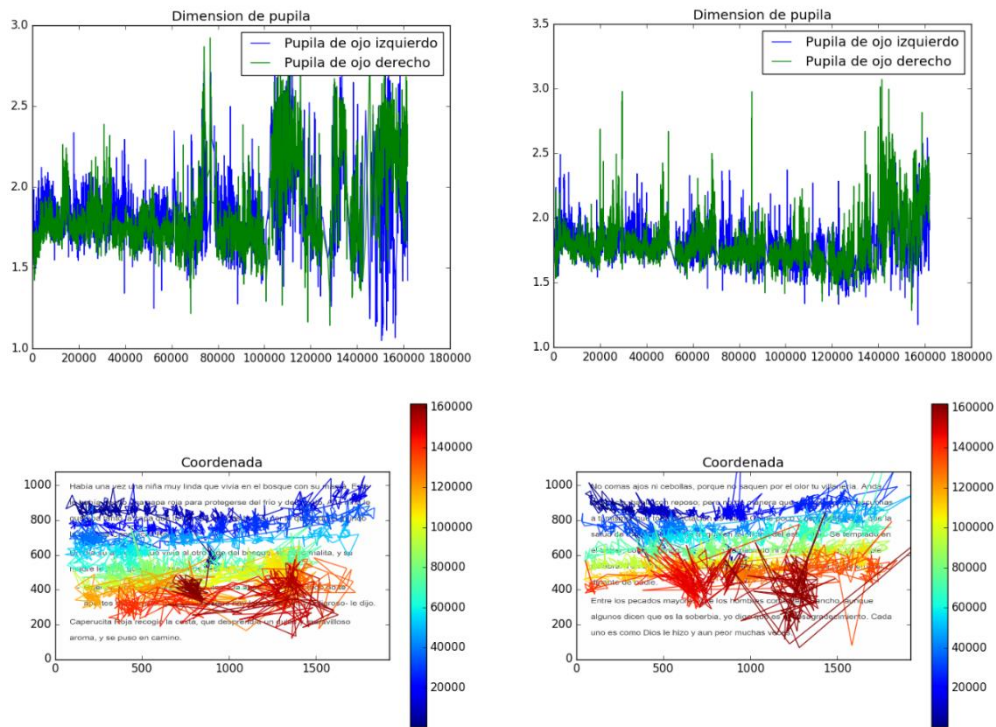


Figura 31. Comparativa de los resultados obtenidos a partir de los datos de seguimiento ocular en la prueba de lectura de dos textos del paciente M_J. En el lado izquierdo, tenemos la dimensión de la pupila (arriba) y el gráfico de visualización del seguimiento ocular (abajo) para el texto de Caperucita Roja y, en el lado derecho, las mismas imágenes para el fragmento de Don Quijote.

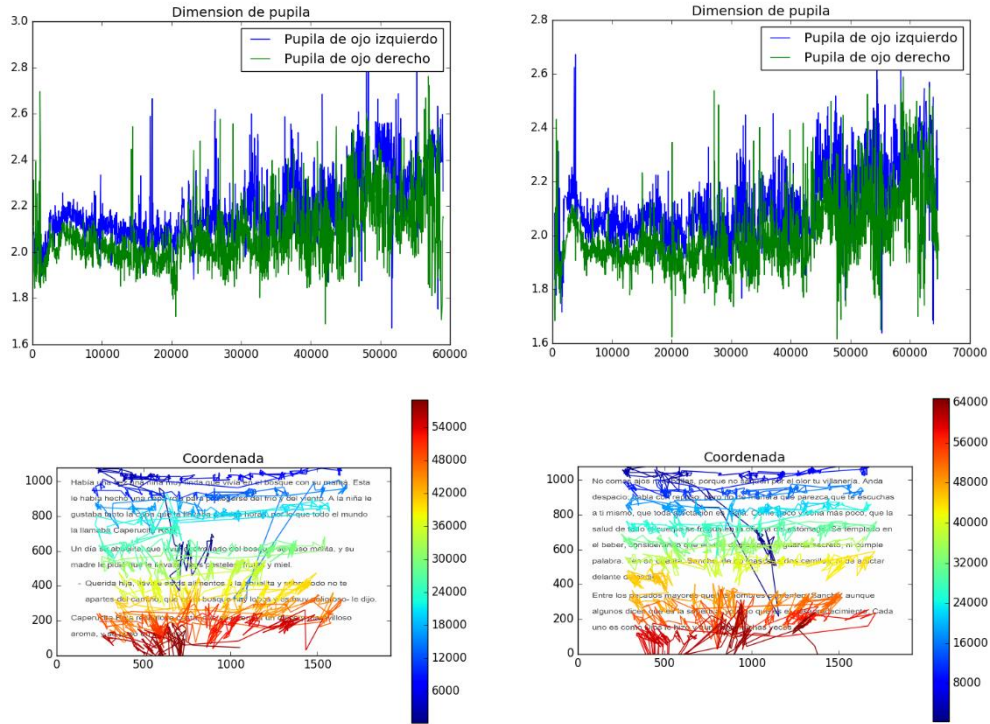


Figura 32. Comparativa de los resultados obtenidos a partir de los datos de seguimiento ocular en la prueba de lectura de dos textos del paciente M_JJ. En el lado izquierdo, tenemos la dimensión de la pupila (arriba) y el gráfico de visualización del seguimiento ocular (abajo) para el texto de Caperucita Roja y, en el lado derecho, las mismas imágenes para el fragmento de Don Quijote.

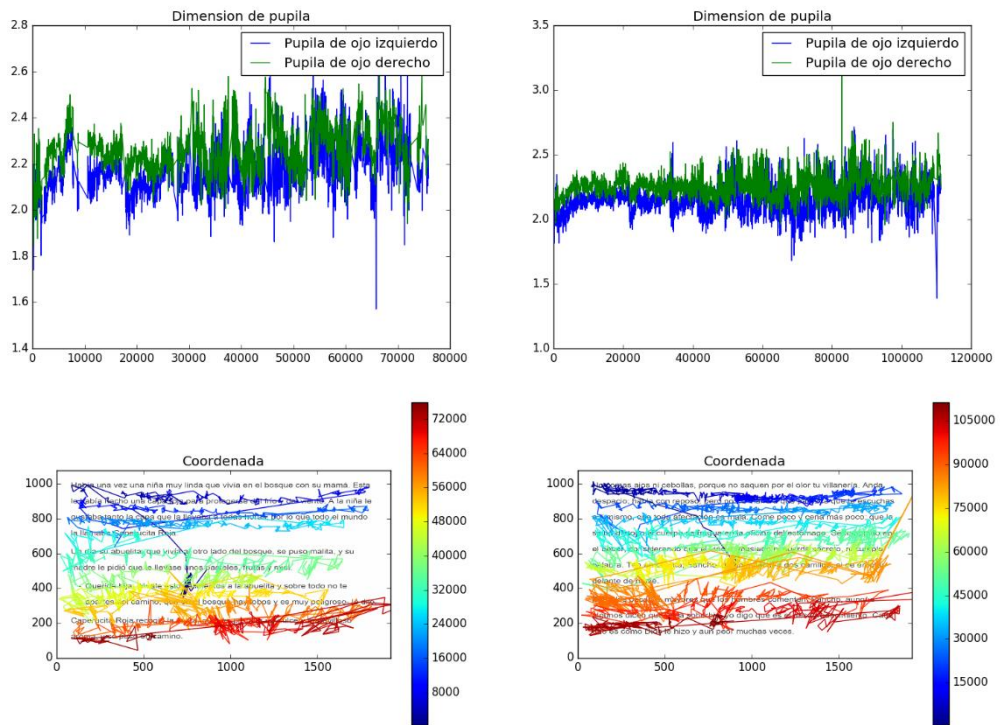


Figura 33. Comparativa de los resultados obtenidos a partir de los datos de seguimiento ocular en la prueba de lectura de dos textos del paciente P_M. En el lado izquierdo, tenemos la dimensión de la pupila (arriba) y el gráfico de visualización del seguimiento ocular (abajo) para el texto de Caperucita Roja y, en el lado derecho, las mismas imágenes para el fragmento de Don Quijote.

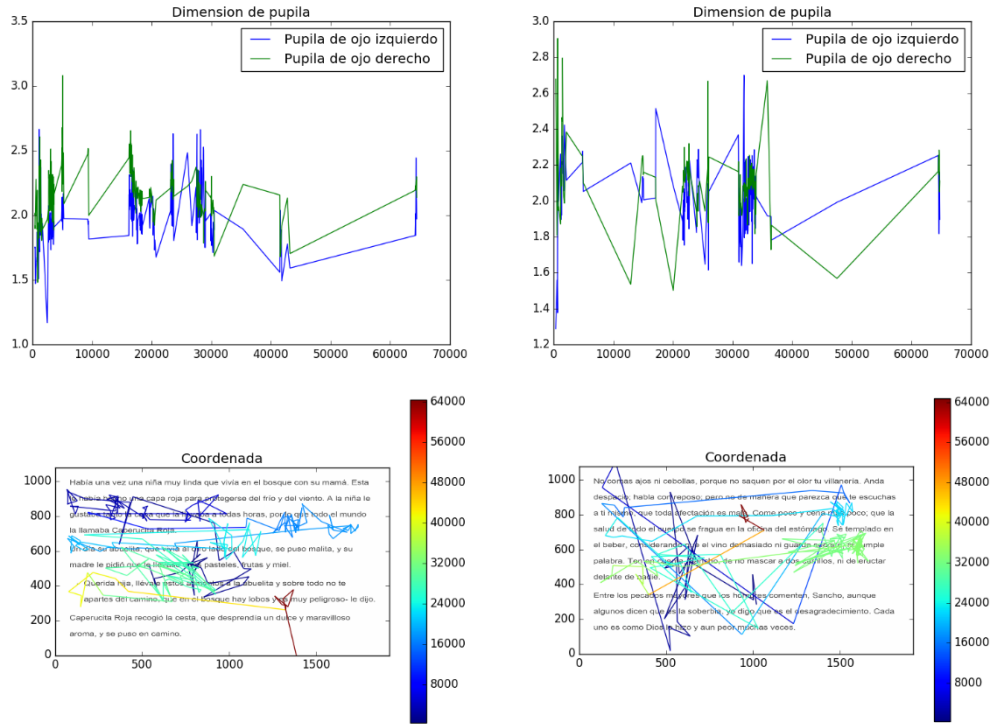


Figura 34. Comparativa de los resultados obtenidos a partir de los datos de seguimiento ocular en la prueba de lectura de dos textos del paciente T_A. En el lado izquierdo, tenemos la dimensión de la pupila (arriba) y el gráfico de visualización del seguimiento ocular (abajo) para el texto de Caperucita Roja y, en el lado derecho, las mismas imágenes para el fragmento de Don Quijote.