

Diagnóstico de TDAH mediante aprendizaje profundo basado en la combinación de técnicas de neuroimagen y características lingüísticas.

Paula García López
paula.garcial@doctorandos.uctv.es
Universidad Complutense de
Torrelodones de la Vega
Murcia, España

Adrián Rubio Pintado
adrian.rubiop@estudiante.uam.es
Universidad Autónoma de Madrid
Madrid, España

Carlos Ramos Mateos
carlos.ramosm@estudiante.uam.es
Universidad Autónoma de Madrid
Madrid, España

Marcos Martínez Jiménez
marcos.martinezj@estudiante.uam.es
Universidad Autónoma de Madrid
Madrid, España

Alberto Altozano Fernández
alberto.altozano@estudiante.uam.es
Universidad Autónoma de Madrid
Madrid, España

Santiago Montoso Fernández
csantiago.montoso@estudiante.uam.es
Universidad Autónoma de Madrid
Madrid, España

ABSTRACT

El trastorno por déficit de atención e hiperactividad (TDAH) es una de las enfermedades neurológicas más prevalentes en la infancia y la adolescencia. El diagnóstico tradicional es un proceso lento y costoso que sufre de subjetividad y sesgos demográficos, lo que ha suscitado un gran interés en alternativas de aprendizaje automático. En los últimos años se han utilizado técnicas de neuroimagen para diagnosticar diferentes tipos de enfermedades neurológicas, incluido el TDAH, aunque todavía no se ha conseguido alcanzar a los métodos convencionales. En algunas enfermedades también se han aplicado técnicas de procesamiento de lenguaje natural (PLN) para extraer información lingüística de los pacientes, aunque todavía no se ha hecho para el TDAH. En este artículo se propone un modelo de aprendizaje profundo para diagnosticar TDAH en base a información de neuroimagen y características lingüísticas. Para ello se desarrolla un modelo independiente por cada fuente de datos y un modelo final que combina los resultados de los anteriores. También se presenta un nuevo conjunto de datos estándar desarrollado expresamente para la clasificación mediante información combinada. La evaluación de los modelos desarrollados sobre este conjunto muestra una mejora significativa frente al estado del arte que permite plantearlos por primera vez como una alternativa diagnóstica convencional. Además, tanto el conjunto de datos como los modelos han sido desarrollados para minimizar los sesgos del diagnóstico, mitigando así uno de sus principales conflictos éticos.

KEYWORDS

TDAH, trastorno de desarrollo neurológico, EEG, MRI, fMRI, aprendizaje automático, aprendizaje profundo, diagnóstico, CNN, LSTM, PLN, BERT

ACM Reference Format:

Paula García López, Adrián Rubio Pintado, Carlos Ramos Mateos, Marcos Martínez Jiménez, Alberto Altozano Fernández, and Santiago Montoso Fernández. 2018. Diagnóstico de TDAH mediante aprendizaje profundo basado en la combinación de técnicas de neuroimagen y características lingüísticas. In *Proceedings of* . ACM, New York, NY, USA, 24 pages. <https://doi.org/10.1145/nnnnnnnn.nnnnnnnn>

1 INTRODUCCIÓN

1.1 Trastorno por déficit de atención e hiperactividad

El trastorno por déficit de atención e hiperactividad (TDAH), es una de las afecciones neurológicas con mayor incidencia entre los niños, afectando a alrededor del 4% de la población menor de 12 años. Sus síntomas principales incluyen, entre otros, la distracción, inquietud, falta de concentración, hiperactividad, irritabilidad e impulsividad. [16] Estos síntomas tienen graves consecuencias sobre la socialización y el desarrollo académico durante la infancia [78], y a menudo persisten hasta la edad adulta, donde dificultan el mantenimiento de las relaciones personales y pueden reducir drásticamente la productividad del individuo [5].

El TDAH también tiene una importante dimensión social que afecta al diagnóstico. Se ha demostrado que el subtipo de TDAH hiperactivo (TDAH II) se manifiesta principalmente en varones, mientras que el subtipo no hiperactivo (TDAH I) está asociado a pacientes femeninos [46, 70]. El comportamiento hiperactivo es una de las señales más evidentes del trastorno, por lo que esto da lugar a que se diagnostique con mucha menor frecuencia en mujeres. Algunos autores han propuesto que no hay una diferencia real entre ambos subtipos, sino que son el resultado de diferencias sociales en la educación de niños y niñas [70] que favorecen la interiorización de los síntomas. Pese a que la tasa de mujeres entre los pacientes diagnosticados con TDAH está en torno al 10-20%, estudios que tienen en cuenta las diferencias de género sugieren que la enfermedad de hecho tiene una incidencia similar entre mujeres y hombres [40, 81].

Sin embargo, incluso en el mejor de los casos el diagnóstico convencional de TDAH es ampliamente subjetivo y está plagado

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than ACM must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from permissions@acm.org.

© 2018 Association for Computing Machinery.
<https://doi.org/10.1145/nnnnnnnn.nnnnnnnn>

de errores. Estudios con cientos de miles de pacientes revelan altas tasas de diagnóstico incorrecto [27] y hasta un 20% de falsos positivos [70].

Es por esto que existe un gran interés en el campo sanitario por desarrollar nuevos métodos de diagnóstico que permitan la detección temprana y libre de sesgos del TDAH.

1.2 Técnicas de diagnóstico actuales

Los métodos de diagnóstico más usados actualmente consisten en entrevistas, cuestionarios e inspecciones por parte de psiquiatras infantiles. Éstos son marcadamente subjetivos y suelen dar lugar a resultados poco fiables o confusión con otras patologías, lo que ha suscitado gran interés en el desarrollo de técnicas diagnósticas basadas en información objetiva como datos personales de los pacientes o análisis de neuroimagen.

En los últimos años ha crecido ampliamente el uso de técnicas de neuroimagen médica, principalmente electroencefalografía (EEG), tomografía por resonancia magnética (MRI) y derivados como el MRI funcional (fMRI), para el diagnóstico de distintas afecciones neurológicas [6]. Estas técnicas revelan aspectos morfológicos y funcionales del cerebro con alta resolución y son de especial interés en patologías donde la causa biológica sigue siendo mayormente desconocida, como la enfermedad de Alzheimer (EA), el trastorno del espectro autista (TEA), o el propio TDAH.

La señal EEG es barata y accesible, pero solo mide la actividad global en una serie de zonas superficiales del cerebro (canales). En cambio, las señales fMRI miden la actividad en cada punto tridimensional del cerebro, pero son más costosas y están a disposición de solo unos pocos hospitales. Pese a su interés en enfermedades con causa biológica desconocida, el propio desconocimiento de la enfermedad ha impedido encontrar reglas de diagnóstico fiables a partir de información clínica o de neuroimagen, lo que justifica el reciente interés en desarrollar métodos de diagnóstico basados en aprendizaje automático que puedan extraer la información discriminativa contenida en las señales de EEG o fMRI de forma más efectiva que los análisis manuales.

1.3 Diagnóstico por aprendizaje automático

Las técnicas de aprendizaje automático, especialmente con el desarrollo del aprendizaje profundo, han demostrado ser muy útiles para resolver problemas difíciles de abordar para el ser humano en multitud de campos, incluido el diagnóstico médico. Así, trabajos recientes las han aplicado para tratar de predecir ataques epilépticos [73], depresión [39] o EA [45] a partir de señales de EEG o fMRI.

En caso del diagnóstico de TDAH, las señales de EEG se han analizado fundamentalmente mediante modelos clásicos con etapas de preprocesamiento sencillas. En varios trabajos [24, 63, 64] la señal de 2 dimensiones (tiempo y canal) se pasa a 3 dimensiones (tiempo, canal y frecuencia) mediante la obtención de un espectrograma por ventanas temporales. Otra opción frecuente es utilizar características diseñadas manualmente como regiones de interés, zonas del cerebro que se consideran informativas, o la densidad de materia gris [8, 38]. Los modelos de clasificación principales serían máquinas de vectores de soporte (SVMs) [7, 38, 63], aunque hay algunos artículos que aplican aprendizaje profundo [9, 84]. Además, la mayoría de trabajos utilizan conjuntos de datos propios con muy

pocos pacientes (los más grandes no superan los 300 y la mayoría tiene menos de 50) por lo que obtienen métricas de efectividad muy altas pero poco fiables.

Lo contrario ocurre en el caso del fMRI, donde los principales modelos emplean redes profundas, con mínimo preprocesamiento más allá de las etapas estándar de normalización y reducción de ruido, y prácticamente todos los artículos utilizan el conjunto estándar ADHD-200, que cuenta con 937 participantes procedentes de 8 sitios [27]. Las redes más populares son las convolucionales (CNNs) y los *autoencoders* (AEs) [56, 72, 85], pero también hay bastantes arquitecturas híbridas, como combinaciones de CNNs y redes neuronales recurrentes (RNNs) [47, 58] o AEs conectados a otras redes [30]. Algunos trabajos entrenan un modelo diferente para cada sitio (unos 100 o 200 pacientes) y reportan medias de las métricas de efectividad por sitio, lo que da lugar a valores muy altos pero poco fiables porque los tamaños de muestra efectivos son muy pequeños. Otros, en cambio, entrenan y evalúan el mismo modelo en todos los sitios a la vez, lo que da lugar a métricas más moderadas, como un 70% de precisión, donde cabe un gran margen de mejora.

1.4 Diagnóstico por características lingüísticas

Una opción que se ha explorado en distintos artículos para alcanzar dicha mejora sería la incorporación de características adicionales, basadas, por ejemplo, en otras características personales o médicas de los pacientes [2, 34]. En este trabajo se propone otra posibilidad, el uso de características del lenguaje a través de procesamiento de lenguaje natural (PLN).

Las técnicas de PLN han empezado a aplicarse recientemente al diagnóstico médico mediante análisis de historiales clínicos [44, 48, 71, 94], monitorización en redes sociales [62] o características del lenguaje como base para pruebas diagnósticas. Esta última posibilidad ha sido explorada principalmente en la detección de EA, donde se busca identificar señales tempranas de deterioro cognitivo leve [32, 35, 37], aunque también hay abundantes trabajos en TEA [36] y depresión [41], enfermedades donde se producen cambios en las interacciones sociales o habilidades necesarias para el lenguaje, y donde la causa biológica es desconocida. Este también sería el caso del TDAH, donde se ha observado que los síntomas de inatención o impulsividad afectan a diversas características en el ámbito de la interacción social y la pragmática [15, 18].

Pese a la abrumadora evidencia de manifestaciones lingüísticas del TDAH, aún no se ha explorado la posibilidad de incluir características PLN para su diagnóstico. Sin embargo, es razonable suponer que los avances que se han producido en EA y TEA se puedan transferir a esta enfermedad. Un aspecto clave a la hora de incorporar características de PLN es definir el tipo de textos que se van a utilizar. La mayoría de artículos actuales recurren a someter al paciente a pruebas como cuestionarios [59, 66], relatar una historia [12, 52] o mantener una conversación [67], siendo la más frecuente la *cookie theft task* [4, 12, 33, 49]. El otro aspecto clave para obtener buenos resultados es el modelo y preprocesamiento utilizados. En este sentido la bibliografía muestra una gran diversidad, con etapas de preprocesamiento como eliminación de pausas, codificación mediante TF-IDF [4], *word embeddings* [57] o selección de características [26]. Asimismo encontramos clasificadores SVM

[89], CNNs y RNNs [4, 82], CNNs combinadas con redes *long-short term memory* (LSTMs) [47] o BERT[68] (entre otros).

1.5 Aportación y objetivos

En este trabajo se busca dar un primer paso hacia un método de diagnóstico automático de TDAH que iguale o supere los métodos convencionales para permitir la identificación temprana de la enfermedad. Para ello se aborda la predicción TDAH mediante las últimas técnicas de aprendizaje profundo y la combinación de señales de EEG y fMRI junto a características lingüísticas obtenidas por PLN.

A este fin se ha llevado a cabo una revisión exhaustiva del estado del arte en redes neuronales, predicción automática de TDAH y diagnóstico de enfermedades neurológicas por PLN, y se ha desarrollado y publicado en abierto un conjunto de datos con sesgos controlados y más de 4000 pacientes que incluye información de las tres fuentes de datos.

Las principales aportaciones de este trabajo son:

- Sistemas de predicción de TDAH a partir de EEG, fMRI y PLN que igualan y en algún caso superan el estado del arte.
- Obtención de una mejora mucho mayor mediante la combinación de los tres sistemas de predicción.
- Demostrar que las señales de EEG o fMRI y las características lingüísticas contienen información complementaria, de modo que su combinación es necesaria para conseguir un método de diagnóstico comparable al convencional.
- Conjunto de datos que supera las limitaciones de los conjuntos previos, permite evaluar diagnóstico a partir de fuentes de datos múltiples y se postula como un nuevo estándar en la predicción de TDAH.

Además, la herramienta desarrollada en este trabajo, y versiones ampliadas posteriores, permitirá detectar la enfermedad de forma más rápida y barata, mejorando significativamente el pronóstico de los pacientes y sus familias y reducir el gasto público en sanidad y educación.

El resto del artículo se organiza de la siguiente manera. La Sección 2 describe los detalles de la revisión bibliográfica realizada. La Sección 3 presenta el conjunto de datos utilizado y los algoritmos propuestos. La Sección 4 expone la propuesta de experimentación. En la Sección 5 se discuten las conclusiones extraídas del trabajo y, finalmente, en la Sección 6 se dan algunas ideas sobre posibles avances que se podrán dar en el futuro.

2 ESTADO DEL ARTE

Para la revisión bibliográfica del estado del arte llevada a cabo en este trabajo se ha hecho un extenso estudio bibliográfico de 365 artículos, de los cuales se han revisado 110 meticulosamente. El procedimiento detallado se puede consultar en el anexo 7.1.

Los artículos revisados se pueden dividir en tres categorías según la fuente de datos que utilizan y la enfermedad que estudian, que se detallarán a continuación:

- Diagnóstico de TDAH por EEG.
- Diagnóstico de TDAH por fMRI.
- Diagnóstico de otras enfermedades neurológicas por PLN (ya que PLN todavía no se ha aplicado a TDAH).

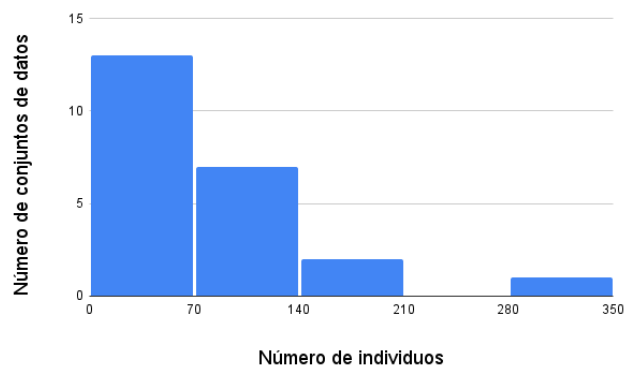


Figura 1: Distribución del número de pacientes en los conjuntos de datos de EEG.

2.1 Diagnóstico de TDAH por EEG

Todos los trabajos de EEG revisados utilizan conjuntos de datos propios, es decir, desarrollados por el grupo de investigación específicamente para estos trabajos. Esto da lugar a un número reducido de pacientes (Figura 1) y mínima reproducibilidad, ya que casi ninguno de ellos es público. La mayoría de conjuntos no alcanzan los 150 pacientes [6, 20, 22, 24, 63] mientras que el más alto solo cuenta con 300 [64]. La actividad realizada durante la toma de datos es diversa, pero en la mayoría se trata de reposo [6, 22, 24, 63, 64] o realizar una tarea sencilla como pulsar un botón inmediatamente tras un estímulo [8, 20, 38, 64].

Respecto al preprocesamiento y selección de las características finales utilizadas, se encuentra que hay bastante diversidad. Algunas etapas, como normalización de la señal, corrección de ruido o filtrado de frecuencias, parecen pasos estándar que se usan en todos los artículos. Posteriormente, en la mayoría se obtiene un espectrograma tridimensional en el dominio de la frecuencia a partir de la señal bidimensional temporal mediante técnicas de enventanado [3, 24, 43, 63, 64]. En otros se aplican técnicas de selección de características para trabajar directamente con un grupo de canales en el dominio de tiempo [3, 6, 8, 13, 91] o se aplica el clasificador directamente [7, 38, 84, 90]. Además, es relativamente frecuente el uso de características no lineales estándar, como la dimensión fractal o medidas de teoría de la información [3, 23, 38, 55].

Los modelos utilizados (ilustrados en la Figura 4) se pueden dividir en tres categorías principales: 1) **clasificadores no neuronales**, 2) **CNNs**(redes convolucionales) y 3) **RNNs (redes recurrentes)**, gran parte de ellas **LSTMs** (redes recurrentes con "memoria"). La gran mayoría de artículos de diagnóstico de TDAH utilizan clasificadores no neuronales, casi todos SVMs [7, 22, 38]. En muchos de ellos se refieren a estos modelos como modelos de aprendizaje profundo, a pesar de no serlo. El principal problema de estos clasificadores es que el estado del arte ya ha demostrado con creces que las redes profundas obtienen resultados mejores consistentemente, especialmente en el campo de análisis de imagen.

Las otras dos categorías corresponderían a redes de tipo CNN [6, 9, 23, 64, 84] y de tipo RNN o LSTM [20, 29, 55, 65]. Las CNNs son algo más frecuentes y, aunque ambos tipos han dado lugar a

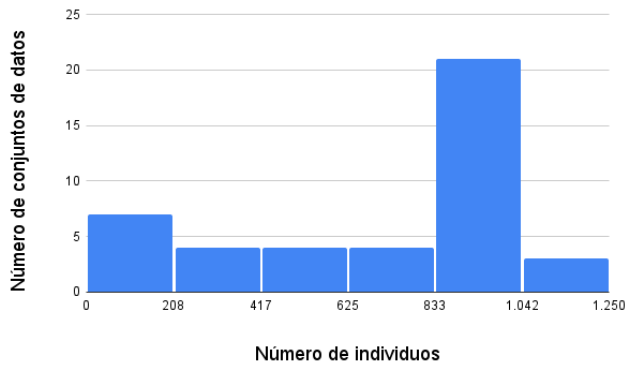


Figura 2: Distribución del número de pacientes en los conjuntos de datos de fMRI.

excelentes resultados, con precisiones de más del 90%, las CNNs los han alcanzado en conjuntos de mayor tamaño. Aunque cabe esperar que una combinación de ambas redes de lugar a resultados aún mejores, aún no se han aplicado este tipo de arquitecturas híbridas en EEG.

Finalmente, respecto a las métricas de evaluación, la precisión sería la métrica por excelencia que usan todos los artículos. Algunos también miden sensibilidad, especificidad o área bajo la curva ROC, pero su utilidad es muy limitada ya que no se pueden comparar con los artículos que solo miden precisión. **Los valores de precisión alcanzados se distribuyen en un grupo con más de 90%** [7–9, 13, 38], donde se llega hasta más de 98% en el mejor de los casos, y otro grupo que oscila entorno a 80% [29, 64, 84, 90]. Además, no parece haber una evolución de la precisión con los años (desde 2017 a 2022) ni con el tamaño del conjunto de datos.

El tamaño reducido de los conjuntos de datos EEG dificulta en gran medida la identificación de modelos adecuados para el diagnóstico. Lo más probable es que las eficacias de más del 90% se deban al pequeño número de pacientes, a la existencia de sesgos o a defectos metodológicos en el entrenamiento y evaluación. Sin embargo, cabe suponer que los mejores modelos utilizados en fMRI, que son mucho más fáciles de identificar debido al gran número de pacientes, también sean los mejores para EEG.

2.2 Diagnóstico de TDAH por fMRI

A diferencia de lo que ocurre en EEG, la mayoría de, los artículos salvo contadas excepciones [31, 60, 83, 88], utilizan el conjunto de datos estándar ADHD-200, que cuenta con 937 participantes provenientes de 8 sitios diferentes [27]. Aunque no todos los artículos consideran los 937 participantes, el resultado son tamaños muestrales mucho mayores (Figura 2). Aun así, cabe mencionar que el conjunto ADHD-200 sigue presentando problemas de sesgo al tener cierto desequilibrio de género y no incluir pacientes con otras enfermedades neurológicas. Por último, algunos artículos también incluyen pacientes con otras enfermedades de otros conjuntos como el ABIDE [19], dando lugar a tamaños aún mayores.

Las etapas de preprocesamiento en trabajos de fMRI también suelen incluir pasos de preprocesamiento estándar de normalización

de la señal, corrección de ruido o filtrado de frecuencias. A continuación, algunos artículos vuelven a utilizar características estándar no lineales [28, 93, 95] y técnicas de selección de características [25, 48, 61, 86, 92]. Sin embargo, aquí es mucho más frecuente obtener otras representaciones de la señal como un grafo de conectividad funcional entre puntos del cerebro (a menudo ROIs elegidas por expertos o de acuerdo a un atlas) [31, 34, 42, 74, 77]. Finalmente, la aplicación directa del clasificador también es más frecuente en este caso, correspondiendo a menudo a un proceso de aprendizaje profundo extremo a extremo [54, 56, 58, 80, 87].

Respecto a la arquitectura de los modelos, prácticamente todos los artículos con fMRI utilizan modelos de aprendizaje profundo (ver Figura 4). Entre éstos se encuentran redes simples como CNNs [11, 17, 28, 56, 95], RNNs [50] o AEs(AutoEncoders) seguidos de otro clasificador [11, 77], pero también abundantes arquitectura híbridas como combinaciones de CNNs con RNNs [1, 10, 58, 80] o AEs seguidos de otras redes [30, 80]. Entre las CNNs, que vuelven a ser las redes más populares usadas para clasificación de imágenes fMRI, encontramos diversas variantes como las redes convolucionales para grafos [87, 93] o las redes convolucionales anidadas [42]. Los AEs también incluyen variantes especiales como los AEs convolucionales [11, 34, 48, 80] o espaciotemporales [30, 72]. Las redes convolucionales, incluyendo las propias CNNs y también los AEs convolucionales, vuelven a ser algunos de los modelos más populares y exitosos. Aunque tamaños muestrales algo menores pueden haber dado lugar a precisiones sobre estimadas.

Las métricas con las que se evalúan los resultados son similares a las de EEG. Todos los artículos miden la precisión de los modelos, mientras que solo algunos miden la sensibilidad, la especificidad o el área bajo la curva ROC. Así, la **precisión** es la única medida que permite comparar los resultados de todos los artículos. Sus valores también se distribuyen en **dos grupos**, uno que supera el **90%** [21, 42, 60, 72, 77], que corresponde a los artículos que entrenan el modelo con los pacientes de cada sitio por separado, y otro que abarca **entre el 70 y el 80%** [34, 48, 50, 56, 79] y que corresponde a los artículos que entrenan el modelo con el conjunto de datos completo. Por tanto, en este caso la precisión si depende del tamaño del conjunto de datos y también parece incrementar ligeramente con los años.

Aunque identificar el mejor modelo posible en el caso de fMRI sigue siendo complicado debido a la variabilidad del origen de los datos y a si se toman juntos o por separado durante el entrenamiento, es más sencillo que en el caso del EEG. Como se mencionaba anteriormente, prácticamente todos los artículos con precisiones del 90% consideran el origen de los datos del ADHD-200 por separado, dando lugar a tamaños muestrales similares a los de EEG, y, por tanto, a precisiones sobre estimadas. En cambio, los artículos que consideran todas las fuentes de datos a la vez dan lugar a precisiones más moderadas y fiables, permitiendo identificar cómo las mejores soluciones las propuestas por [34, 42, 51, 56].

2.3 Diagnóstico de enfermedades neurológicas por PLN

Tal y como se menciona en la Subsección 1.4, todavía no hay artículos que utilicen características de PLN para diagnosticar TDAH. De

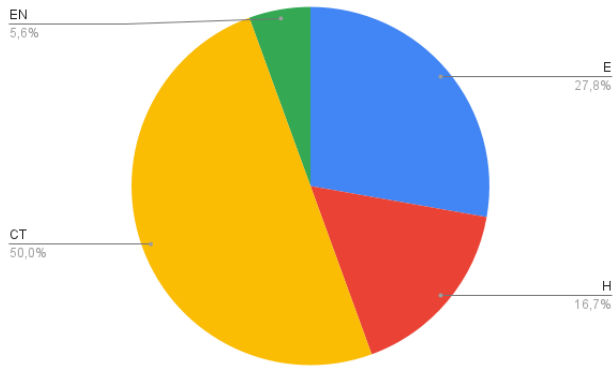


Figura 3: Distribución de la actividad utilizada en los conjuntos de datos de PLN. Se muestra el porcentaje de conjuntos que utilizan entrevistas (E), la tarea *Cookie Theft* (CT), enumeración de una lista (EN), contar una historia (H), discursos (D) y actividad cotidiana (AC).

hecho, aunque se utilizan técnicas de PLN, como análisis de historiales médicos, para diagnosticar otras enfermedades, las únicas para las que se han utilizado características lingüísticas son la EA, con un 85% de los artículos, y el TEA, con el 15% restante.

Aunque en este caso el 63% de los trabajos utiliza conjuntos de datos estándar, como el ADReSSo[76] o el DementiaBank [14], la distribución de tamaños es relativamente similar a la de EEG. El conjunto más grande solo incluye 310 pacientes [47] y la mitad tiene menos de 125 [26, 37, 49, 52, 82]. A diferencia de EEG y fMRI, la actividad realizada durante la toma de datos es crítica para PLN, ya que determina las características lingüísticas que pueden ser detectadas. La Figura 3 muestra la distribución de actividades en el estado del arte. Las actividades más populares son la tarea *Cookie Theft* [4, 26, 33, 53, 89], un estándar para muchas enfermedades, y las entrevistas [37, 47, 57, 67, 69].

Respecto al preprocesamiento, algunos artículos utilizan un esquema muy específico, como [68], que elimina las pausas, aplica un paso de *POS tagging* para extraer sustantivos y adjetivos de los que obtiene una representación por *word embeddings*, utiliza los *word embeddings* para agrupar las palabras por KMeans, y finalmente clasifica en base al grupo asignado. Sin embargo, la gran mayoría de artículos aplican unos pocos pasos, como eliminación de signos, puntuación y espacios [4, 32, 49, 53, 89], codificación del texto mediante TF-IDF [4, 49, 52, 89] o *word embeddings* [32, 52, 53, 57, 89] o selección de características [26, 53, 82, 89]. En otros trabajos se utiliza un preprocesamiento mínimo al utilizar una red profunda como BERT directamente sobre la entrada [53, 69, 75].

Entre los modelos utilizados (ver Figura 4) encontramos una distribución relativamente homogénea entre modelos clásicos (SVMs [26, 37, 67, 68, 89], Random Forest o Naive Bayes [12]), RNNs [4, 33, 47, 52, 82], CNNs [4, 47, 57] y transformers como BERT [49, 53, 68, 69, 75]. También encontramos algunos casos de CNNs combinadas con RNNs [4, 47].

Al igual que en los casos anteriores, cada artículo mide métricas diferentes por lo que la **precisión**, utilizada en casi todos, acaba

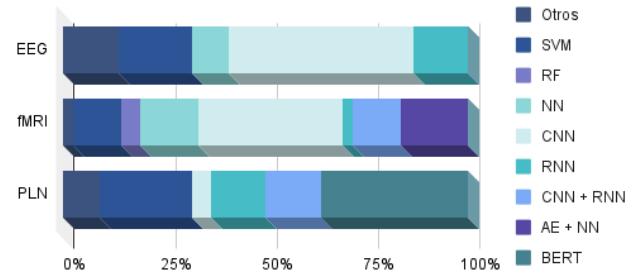


Figura 4: Distribución de los modelos aplicados en cada fuente de datos. Las siglas RF se refieren a *Random Forest* y NN a *red neuronal*.

siendo la única que se puede comparar entre artículos. Su valor oscila entre un **80% y casi un 100%**, sin ninguna evolución temporal apreciable. Los modelos que mayor eficacia alcanzan de entre aquellos casos con más de 200 pacientes son BERT [53, 75, 89], LSTMs [33] y CNNs + LSTMs [4, 47] con precisiones del 93%, 86% y 92% respectivamente. Los tamaños de pacientes de PLN tampoco son los ideales para la selección de modelos, pero todas estas arquitecturas son lo último en el estado del arte de PLN, por lo que cabría esperar que den buenos resultados.

2.4 Resumen

La Figura 4 muestra la distribución de modelos utilizados en la bibliografía para cada fuente de datos. Por otro lado, la Tabla 2 resume los artículos revisados para este trabajo, incluyendo aspectos como el tamaño de los conjuntos de datos, los modelos utilizados o la máxima precisión alcanzada.

De acuerdo a lo expuesto en las subsecciones anteriores, hay un grado de solapamiento importante entre los mejores modelos para distintas fuentes de datos. Las CNNs son de los mejores modelos en EEG [9, 84] y fMRI [42, 56, 93]. Al combinarlas con LSTMs se vuelven a conseguir unos de los mejores resultados en fMRI [58] pero también PLN [4, 47]. En fMRI también destaca el rendimiento de algunas versiones de AE seguidas de otra red neuronal [34, 48]. De forma similar, el modelo ganador para PLN sería BERT [53, 75], que puede considerarse un caso de AE seguido de una red neuronal.

En todos los casos el estado del arte ha alcanzado **precisiones de más del 90%, pero esto solo ocurre con números de pacientes muy reducidos**. Cuando se consideran del orden de mil pacientes el rendimiento baja hasta apenas el 80%, y cabe esperar que bajara aún más en condiciones reales, lo que ilustra que todavía existe un gran margen de mejora.

La oportunidad científica del presente trabajo comienza con utilizar PLN como técnica novedosa para el diagnóstico de TDAH. Combinando esta nueva información a las señales EEG y fMRI, para demostrar que esta información está correlacionada y se puede usar para obtener un método de diagnóstico fiable que supere al estado del arte. Usando para estas 3 fuentes de datos 3 modelos independientes (finalmente combinados) basados en las técnicas y variantes deep learning más punteras existentes hasta la fecha, que han demostrado ser capaces de mejorar las métricas de los clasificadores existentes. Todo ello con un dataset creado y por

primera vez suficientemente grande su uso en los algoritmos de deep learning. También siendo fuente de novedad el control de los sesgos en los datos utilizados para el diagnóstico de la enfermedad.

3 METODOLOGÍA

En este artículo se trata de resolver un problema de clasificación entre pacientes con o sin TDAH mediante la combinación de tres fuentes de información, señales de EEG, fMRI y características lingüísticas obtenidas mediante análisis de PLN.

Para ello se desarrollaron modelos de predicción independientes a partir de cada fuente de datos y un sistema posterior para combinar sus resultados y obtener la predicción final. También se presenta el conjunto de datos abierto con información de EEG, fMRI y PLN que se creó para permitir el diagnóstico combinado.

Los modelos independientes que se desarrollaron igualan o superan los mejores resultados del estado del arte, pero todavía son insuficientes para reemplazar o incluso complementar el diagnóstico convencional. El modelo combinado, en cambio, da lugar a una mejora cualitativa que permite postularlo como el primer método de diagnóstico automático apto para ser implementado en el sistema sanitario.

3.1 Recogida de datos

La recogida de datos se realizó siguiendo los consejos y pautas de un comité ético especializado, y de acuerdo a la legalidad española vigente en materia de protección de datos de la salud presente en la Ley Orgánica de Protección de Datos de Carácter Personal. Para ello se solicitó (en los casos en los que fuera necesario) que los tutores legales de los participantes firmaran un documento de cesión y distribución anonimizada de los datos.

El conjunto de datos publicado (ADHD-4k) contiene individuos sanos y pacientes de distintas enfermedades relacionadas con la salud mental. Durante su creación se veló en todo momento por la diversidad de género, edad y otras variables demográficas de los participantes. Para ello se diseñaron campañas de difusión del estudio clínico dirigidas a los grupos demográficos que estaban infrarepresentados en los pacientes reclutados hasta ese momento.

En total, el conjunto de datos cuenta con 4131 participantes a los que se sometió a la misma prueba para medir sus señales de fMRI, EEG y sus características lingüísticas. La Tabla 1, que resume la demografía de los participantes, permite comprobar que el conjunto es balanceado en términos de género, pacientes de TDAH y edad. Además, el conjunto se ha dividido de forma equitativa durante el entrenamiento para minimizar los sesgos en el aprendizaje de los modelos. Dado el gran volumen del conjunto de datos, esto no debería constituir un problema de cara al rendimiento final obtenido.

Los datos de EEG se recogieron utilizando 32 canales, de acuerdo al sistema internacional 10/20 para favorecer la versatilidad de los datos en el ámbito científico. La tasa de muestreo utilizada fue de 128 Hz. A su vez, las medidas de fMRI se realizaron mediante un proceso estándar internacional utilizado en hospitales de todo el mundo.

En cuanto a las características lingüísticas, el conjunto de datos contiene transcripciones de audio de los pacientes mientras resuelven la tarea *Cookie Theft task*. En dicha prueba se presenta al

	Partición Entrenamiento				Partición Test			
	Niños (5-12)	Joven (12-20)	Adulto (20-65)	Mayor (>65)	Niños (5-12)	Joven (12-20)	Adulto (20-65)	Mayor (>65)
H/M (%)	52/48	46/54	47/53	52/48	49/51	47/53	52/48	50/50
TDAH (%)	45	47	46	45	48	46	47	45
OE (%)	17	23	20	19	18	22	21	20

Tabla 1: Demografía de los individuos incluidos en el conjunto ADHD-4k. Se muestra el porcentaje dentro de cada categoría de edad de: hombres (H), mujeres (M), pacientes TDAH (TDAH) y pacientes de otras enfermedades (OE).

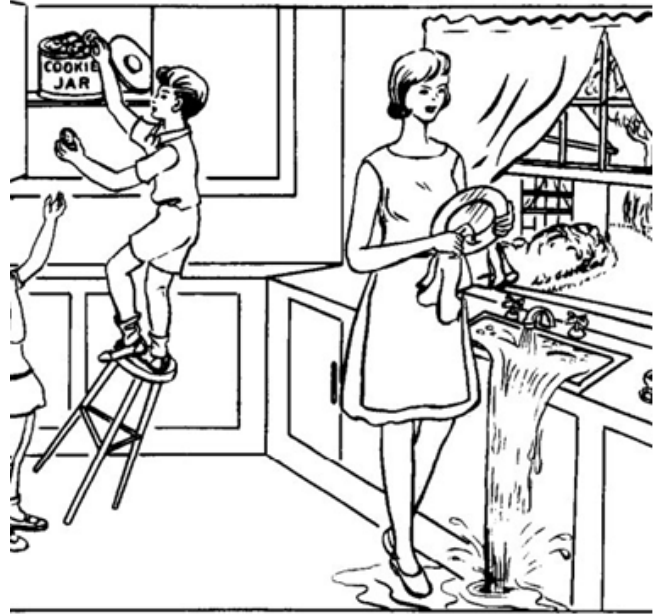


Figura 5: Imagen de la tarea *Cookie Theft*. Se muestra la imagen que los pacientes deben describir durante la tarea propuesta para recoger los datos de PLN.

paciente la imagen de la Figura 5 para que describa qué está sucediendo en ella. Durante el desarrollo de la prueba el entrevistador puede hacer las intervenciones que crea convenientes, ya que el ejercicio está diseñado para simular una conversación sobre un tema único para todos los pacientes. Esta prueba fue elegida bajo el consejo de personal especializado en psicología y procesamiento del lenguaje, ya que permite identificar los problemas de pragmática (tendencia a interrumpir o a divagar, entre otros) que suelen sufrir los pacientes de TDAH.

3.2 Modelo para EEG

A continuación se expone el preprocesamiento y el modelo propuesto en este artículo para la clasificación de los pacientes mediante señales EEG.

Preprocesamiento

En general, las medidas de EEG son muy variables y contienen muchas anomalías, de modo que el preprocesamiento de los datos es clave para un correcto funcionamiento del algoritmo. Las técnicas de preprocesamiento utilizadas son: (1) Reducción de ruido (2) Eliminación de anomalías (3) Normalización Min-Max (4) División en ventanas de 2s.

Debido a la diferente duración de cada una de las muestras, se utiliza el tamaño de muestra más pequeño posible (38 minutos y 40 segundos), lo que da lugar a un total de 1142 ventanas temporales para cada muestra del conjunto de datos.

Modelo propuesto

De acuerdo a lo expuesto en la Sección 2, uno de los modelos que obtienen un mejor rendimiento en fMRI es el *Convolutional Variational Autoencoder* (CVAE). Sin embargo, éste aún no ha sido aplicado a señales de EEG, por lo que en este artículo se plantea una arquitectura similar a la propuesta por *Ming-Shan Gao et al.* [34].

En su artículo, *Ming-Shan Gao et al.*, aplican un CVAE sobre la matriz de correlación de la señal fMRI entre distintas ROIs del cerebro. En el caso de la señal de EEG, esto correspondería a una matriz con las correlaciones entre los 32 canales para cada ventana temporal, dando lugar a un tensor de dimension $32 \times 32 \times 1142$ que se reduce a $32 \times 32 \times 2$ al obtener la media y la varianza a través de las ventanas. La medida de correlación escogida fue la información mutua, que a diferencia de la correlación de Pearson, usada por *Ming-Shan Gao et al.*, permite detectar relaciones no lineales entre canales.

La red CVAE final consta de un codificador y un decodificador simétricos de 4 capas, con *kernels* de Fisher cuyo tamaño será escogido en función del rendimiento del algoritmo. La salida es un vector de 3 componentes que se usa posteriormente para la clasificación y para entrenar el decodificador. Como función de pérdida se utiliza el error de reconstrucción de la imagen original para el autoencoder y el error de predicción para el clasificador.

3.3 Modelo para fMRI

Preprocesamiento

El preprocesamiento de la señal de fMRI se ha llevado a cabo mediante la herramienta fMRIPrep. En primer lugar se elimina el movimiento de las imágenes y se corrige el desfase temporal. Posteriormente se utiliza un filtro de suavizado gaussiano de FWHM = 5mm y se aplica un filtro paso alto para reducir el ruido de la imagen.

Finalmente, al igual que en el caso de EEG, los datos se dividen en ventanas temporales de 2 segundos. En este caso se cuenta con un total de 2395 ventanas por paciente.

Modelo propuesto

El modelo propuesto en este artículo es una red híbrida de aprendizaje profundo que combina una CNN-3D, con la que Abdolmaleki y Saniee [2] alcanzaron uno de los mejores resultados hasta la fecha en clasificación de fMRI, con una LSTM.

Las redes convolucionales están diseñadas para detectar patrones locales en las señales, mientras que las redes LSTMs modelizan secuencias a largo plazo. De esta forma, ambas redes se complementan dando lugar a uno de los modelos más efectivos.

La CNN-3D propuesta por Abdolmaleki y Saniee utiliza la señal de fMRI, que es un tensor de tres dimensiones espaciales y una temporal, reducida a las tres dimensiones espaciales mediante un preprocesamiento que evalúa la actividad general en el tiempo de cada vóxel. Al aplicar una combinación CNN+LSTM, nuestra propuesta permite mantener la información temporal al introducir un tensor espacial de cada ventana secuencialmente. La predicción final se obtiene mediante voto mayoritario de las predicciones obtenidas con las últimas 100 ventanas temporales.

El número y tamaño de capas en la CNN-3D es difícil de establecer, ya que Abdolmaleki y Saniee no dan detalles al respecto y hay gran variabilidad en los que usan otros artículos de CNNs menos exitosos. Debido a ello, la arquitectura básica considerada incluye de 2 a 4 capas convolucionales y de 1 a 3 capas LSTM. La configuración final se elige mediante la selección de los hiperparámetros durante el proceso de entrenamiento (ver Sección 4).

Finalmente, la función de coste utilizada será la entropía cruzada binaria, un estándar aplicado en prácticamente todas las redes neuronales para clasificación. El entrenamiento también incluye regularización por *dropout* y penalización de los pesos de la red para evitar el *overfitting*.

3.4 Modelo para PLN

Preprocesamiento

El paso de preprocesamiento consiste en limpiar las transcripciones de los pacientes mediante eliminación de mayúsculas, signos de puntuación y otros caracteres especiales introducidos por el software de transcripción.

Modelo propuesto

La ausencia de trabajos previos que utilicen técnicas de PLN para diagnosticar TDAH implica una gran dificultad a la hora de elegir el modelo correcto. Pese a ello, cabe esperar que los modelos que producen mejores resultados para otras enfermedades neurológicas (p. ej. Alzheimer) también produzcan buenos resultados para el TDAH.

Se propone usar como modelo de clasificación para información lingüística el *transformer* BERT, que no sólo es uno de los más exitosos en EA o TEA (Sección 2), sino que también ha alcanzado rendimientos excelentes en numerosas tareas de PLN.

BERT es un codificador bidireccional de Transformadores basado en redes neuronales preentrenadas aplicable a múltiples tareas. En este caso, BERT se utiliza para caracterizar las respuestas de los pacientes y posteriormente para clasificarlas mediante una red neuronal profunda, siguiendo el esquema propuesto por *Ziming Liu et al* [53].

3.5 Modelo global

Una vez entrenados los modelos de fMRI, EEG y PLN por separado, se combinan mediante un modelo global para obtener mejor rendimiento y mayor robustez.

El proceso de combinación implica utilizar la probabilidad de padecer TDAH estimada a partir de cada fuente de datos, es decir, un vector de tres dimensiones, como entrada para un modelo de combinación. Los modelos de combinación considerados son:

- (1) Regresión logística: permite medir la significación estadística de cada fuente de datos mediante un test ANOVA. De esta forma se pueden detectar y eliminar fuentes redundantes para abaratar el método de diagnóstico.
- (2) *Random Forest*: alta eficiencia en el entrenamiento y nula tendencia al sobreajuste.
- (3) Red neuronal profunda: modelo muy flexible. Para evitar el sobreajuste se aplica *dropout* y regularización de los vectores de pesos en la función de coste.

4 PROPUESTA EXPERIMENTAL

4.1 Preprocesamiento del conjunto de datos

El primer paso del diseño experimental es aplicar las etapas de preprocesamiento descritas en la Sección 3 a todos los pacientes del conjunto de datos desarrollado.

Se espera que el conjunto de datos final contenga al menos 4000 individuos, llegando a lo sumo a unos 4500. Cada individuo cuenta con pistas del orden de una hora de señal EEG y fMRI, distribuidas entre 1000 y 2000 ventanas temporales, y unas 100 frases transcritas recogidas durante las entrevistas con los pacientes.

La principal ventaja de contar con un conjunto de gran tamaño y sesgos controlados es conseguir un buen rendimiento para todo tipo de pacientes, independientemente de características como su género o edad. Para evaluar los sesgos del método desarrollado se segregan las métricas de efectividad según el género, edad y otras enfermedades que tengan los pacientes. Teniendo en cuenta el volumen de datos esperados se propone una división de cada factor en 4 categorías o menos, que deberían agrupar suficientes individuos para ser estadísticamente representativas.

- Edad: Niño (5-12 años), Joven (12-20 años), Adulto (20-65 años) y Mayor (>65 años).
- Género: Hombre y Mujer, pudiéndose añadir una categoría de Otros si al menos 50 participantes la eligen.
- Otras enfermedades: Ninguna, No Neurológicas (que puedan afectar a las señales de neuroimagen o el lenguaje), Neurológicas Diferenciables (con síntomas fácilmente distinguibles del TDAH, como el trastorno obsesivo-compulsivo) y Neurológicas Confundibles (con síntomas difíciles de distinguir o que suelen aparecer junto al TDAH.)

Finalmente, el conjunto completo se divide en una partición de entrenamiento y test con una proporción de individuos del 70% y 30% respectivamente. Las características demográficas de las particiones de entrenamiento y test se resumen en la Tabla 1.

4.2 Entrenamiento de los modelos

La partición de entrenamiento se divide en otras 10 particiones de selección de hiperparámetros y validación mediante validación cruzada.

Sobre cada partición de selección se entrenan por separado los modelos propuestos con cada fuente de datos. Para cada fuente de

datos se entrenan 10 variaciones del modelo propuesto con distinta arquitectura, variando los siguientes hiperparámetros:

- EEG: número de capas del CVAE y dimensión del vector codificado.
- fMRI: número de pares de capas convolucional-*pooling* y número de capas LSTM.
- PLN: tasa y tamaño de lotes usados en el aprendizaje.

Así mismo, se entrenan los tres modelos descritos en la Sección 3 para obtener los tres posibles modelos combinados.

Una vez entrenada cada versión del modelo de cada fuente de datos y los tres modelos combinados, se utiliza la partición de validación correspondiente para estimar su precisión y sensibilidad dentro de las categorías demográficas establecidas en la subsección anterior. La media armónica entre las medidas para cada categoría de un factor demográfico (género, edad y enfermedades) se usa para resumir la precisión y sensibilidad del modelo respecto a ese factor. La media armónica aumenta con el aumento de cada valor involucrado en el cálculo, pero penaliza que un valor aumente mucho a costa de otro, por lo que se usa para valorar la distribución igualitaria de los aciertos entre las categorías. De forma similar, la media armónica entre las medidas de cada factor se usa para resumir la efectividad global del modelo. Este es el criterio usado para elegir la mejor arquitectura y el mejor modelo de combinación dentro de una partición de selección y validación.

Los modelos finales para cada fuente de datos y el modelo de combinación final serán los que hayan obtenido mayor promedio de precisión y sensibilidad a través de las 10 particiones. Finalmente, éstos se reentrenan sobre la partición de entrenamiento completa.

4.3 Evaluación

El modelo final de cada fuente de datos se usa para diagnosticar el TDAH en los pacientes de la partición de test. Los resultados obtenidos también se usan para obtener la predicción del modelo combinado al aplicar el esquema de combinación final.

La salida de los tres modelos aislados y la del modelo combinado se compararan con las etiquetas reales de los pacientes (sin dividir por categorías) para calcular su precisión, sensibilidad, media armónica entre ambas y el área bajo la curva ROC.

También se obtiene la precisión y sensibilidad agregadas entre los tres factores demográficos y sus categorías mediante la media armónica (al igual que en el entrenamiento), que permiten identificar sesgos en el algoritmo final. A modo de ejemplo, una precisión global alta, pero armónica baja, permite identificar un sesgo en el modelo causado por un exceso de error en uno de los factores demográficos, como el género. Las precisiones para los distintos géneros permiten identificar en última instancia el grupo problemático.

4.4 Comparación con el estado del arte

Se realiza un proceso similar al descrito anteriormente, selección de hiperparámetros, entrenamiento, evaluación por categorías y evaluación global, para las mejores propuestas del estado del arte (descritas al final de la Sección 2) sobre el conjunto de datos desarrollado. Esto permite comparar los modelos en igualdad de condiciones, es decir, bajo el mismo conjunto, con hiperparámetros seleccionados de acuerdo al conjunto y con las mismas métricas de rendimiento.

Se espera que, al menos, la propuesta supere a las de la bibliografía en términos de precisión y sensibilidad armónica. Todos los trabajos hasta la fecha trabajan para maximizar el acierto general de sus algoritmos sin tener en cuenta la distribución del acierto sobre grupos demográficos, por lo que el modelo propuesto, que sí promueve un tratamiento igualitario, debe mejorar los resultados.

Así mismo, el modelo utilizado y la combinación de tres fuentes de datos, muy diversas, también justifica esperar mejores resultados en términos de acierto general. Las arquitecturas de CVAE y CNN+LSTM y utilizadas son la punta de lanza del aprendizaje profundo para clasificación de señales, y las que mejor han funcionado en clasificación de TDAH. Algo similar ocurre con BERT en el campo de diagnóstico por PLN. Por otro lado, es de esperar que las señales de neuroimagen y las características lingüísticas contengan información complementaria que mejore la clasificación en los casos más complejos o en presencia de enfermedades confundibles con TDAH.

5 CONCLUSIONES

En este trabajo se ofrece una nueva propuesta para el diagnóstico automático de TDAH, que mejora los métodos de diagnóstico subjetivos actuales. La arquitectura se divide en 3 modelos deep-learning para las 3 diferentes fuentes empleadas: un CVAE(Convolucional Variational Autoencoder) para las señales EEG, un modelo híbrido CNN-3D + LSTM para las señales fMRI, y finalmente BERT (un codificador bidireccional de *transformers* basado en redes neuronales preentrenadas) para las características lingüísticas que se incorporan como fuente novedosa de datos, no utilizada anteriormente en el estado del arte para el diagnóstico de la enfermedad, a pesar de que la evidencia médica afirma ser un síntoma determinante de la misma.

Los datos utilizados para entrenar y validar los modelos pertenecen a un dataset de creación propia: *ADHD-4K*, con un tamaño de 4.000 individuos, siendo el dataset de mayor tamaño y libre de sesgos utilizado hasta la fecha para el estudio del TDAH.

Utilizando validación cruzada con 10 pliegues, se escogen los 3 mejores modelos y se combinan sus salidas para así obtener finalmente en un clasificador para el diagnóstico muy robusto.

La herramienta desarrollada en este trabajo permite detectar la enfermedad TDAH de forma rápida y barata, pudiéndose tomar medidas de tratamiento tempranas. Así, al disponer de datos con sesgos controlados, se puede detectar la presencia de la enfermedad en pacientes con síntomas leves y que no se diagnostican. De este modo, se reduce el impacto negativo del TDAH en los pacientes y sus familias. Esto se traduce en una reducción del gasto público en sanidad y educación.

La propuesta tiene ciertas limitaciones. Entre ellas se incluye la limitada experiencia en el uso de técnicas PLN para el diagnóstico de enfermedades como TDAH. En algunos casos, es posible que se proporcionen datos mal etiquetados a los algoritmos, aunque se siga un criterio médico experto. Además, a pesar de que el tamaño del dataset creado es muy superior al usado en la literatura, puede que sea necesario un tamaño de al menos 1 o 2 órdenes mayor para poder obtener un sistema de diagnóstico realmente funcional basado en algoritmos de deep learning.

6 TRABAJO FUTURO

Las propuestas de trabajo futuro se pueden englobar en varias líneas de investigación.

Las técnicas basadas en PLN para detección de TDAH son un área de investigación en si mismo ya que hasta la fecha únicamente se han empleado para detectar enfermedades neurodegenerativas como el Alzheimer. Por ello, los métodos diagnósticos utilizados, como el tipo de entrevistas o las descripciones, se deben adaptar y especializar a la enfermedad TDAH. Así mismo, las propias técnicas de procesamiento de lenguaje están en desarrollo y se prevé que en los próximos años arrojen resultados muy prometedores.

Los modelos de aprendizaje profundo son objeto de constante investigación y conforme avance la capacidad de cómputo las técnicas empleadas irán evolucionando a modelos más complejos. Uno de los modelos más prometedores son las redes CNN 4D que incluyen la información temporal y que integran la imagen espacial tridimensional con la coherencia temporal en el aprendizaje de la red. El uso de estas redes debe ir acompañado con la generalización del uso de aparatos fMRI 4D en los centros de diagnóstico.

La necesidad de conjuntos de datos de gran volumen para entrenar adecuadamente las redes profundas, así como la necesidad de disponer de bases de datos universales que permitan el uso de las técnicas diagnósticas por la totalidad de la población, sugiere concentrar parte del esfuerzo de investigación en seguir mejorando los conjuntos de datos, tanto en cantidad como en calidad.

ACKNOWLEDGMENTS

A los hospitales y centros clínicos que colaboraron con nuestro grupo de investigación en el proyecto científico que dio lugar al conjunto ADHD-4k y a este artículo. Gracias a ellos se pudo obtener un conjunto equilibrado y de mayor tamaño hasta la fecha en predicción de TDAH.

REFERENCES

- [1] 2020. Separated Channel Attention Convolutional Neural Network (SC-CNN-Attention) to Identify ADHD in Multi-Site Rs-fMRI Dataset. *Entropy* 22 (8 2020), 893. Issue 8. <https://doi.org/10.3390/e22080893>
- [2] Sahar Abdolmaleki and Mohammad Saniee Abadeh. 2020. Brain MR Image Classification for ADHD Diagnosis Using Deep Neural Networks. *2020 International Conference on Machine Vision and Image Processing (MVIP)*, 1–5. <https://doi.org/10.1109/MVIP49855.2020.9116877>
- [3] Marco Aceves-Fernandez. 2021. Methodology proposal of ADHD classification of children based on cross recurrence plots. *Nonlinear Dynamics* 104 (04 2021). <https://doi.org/10.1007/s11071-021-06336-z>
- [4] Surabhi Adhikari, Surendrabikram Thapa, Usman Naseem, Priyanka Singh, Huan Huo, Gnana Bharathy, and Mukesh Prasad. 2022. Exploiting linguistic information from Nepali transcripts for early detection of Alzheimer's disease using natural language processing and machine learning techniques. *International Journal of Human-Computer Studies* 160 (4 2022), 102761. <https://doi.org/10.1016/j.ijhcs.2021.102761>
- [5] Jessica C. Agnew-Blais, Guilherme V. Polanczyk, Andrea Danese, Jasmin Wertz, Terrie E. Moffitt, and Louise Arseneault. 2018. Young adult mental health and functional outcomes among individuals with remitted, persistent and late-onset ADHD. *The British Journal of Psychiatry* 213 (9 2018), 526–534. Issue 3. <https://doi.org/10.1192/bjp.2018.97>
- [6] Amirasoud Ahmadi, Mehrdad Kashefi, Hassan Shahrokhi, and Mohammad Ali Nazari. 2021. Computer aided diagnosis system using deep convolutional neural networks for ADHD subtypes. *Biomedical Signal Processing and Control* 63 (1 2021), 102227. <https://doi.org/10.1016/j.bspc.2020.102227>
- [7] Alaa Eddin Alchalabi, Shervin Shirmohammadi, Amer Nour Eddin, and Mohamed Elsharnouby. 2018. FOCUS: Detecting ADHD Patients by an EEG-Based Serious Game. *IEEE Transactions on Instrumentation and Measurement* 67 (7 2018), 1512–1520. Issue 7. <https://doi.org/10.1109/TIM.2018.2838158>

- [8] Miray Altınkaynak, Nazan Dolu, Ayşegül Güven, Ferhat Pektaş, Sevgi Özmen, Esra Demirci, and Meltem İzzetoğlu. 2020. Diagnosis of Attention Deficit Hyperactivity Disorder with combined time and frequency features. *Biocybernetics and Biomedical Engineering* 40 (7 2020), 927–937. Issue 3. <https://doi.org/10.1016/j.bbe.2020.04.006>
- [9] Patricia Amado-Caballero, Pablo Casaseca de-la Higuera, Susana Alberola-Lopez, Jesus Maria Andres de Llano, Jose Antonio Lopez Villalobos, Jose Ramon Garmendia-Leiza, and Carlos Alberola-Lopez. 2020. Objective ADHD Diagnosis Using Convolutional Neural Networks Over Daily-Life Activity Records. *IEEE Journal of Biomedical and Health Informatics* 24 (9 2020), 2690–2700. Issue 9. <https://doi.org/10.1109/JBHI.2020.2964072>
- [10] Abhay Aradhya, Aditya Joglekar, Sundaram Suresh, and Mahardhika Pratama. 2019. Deep Transformation Method for Discriminant Analysis of Multi-Channel Resting State fMRI. *Proceedings of the AAAI Conference on Artificial Intelligence* 33 (07 2019), 2556–2563. <https://doi.org/10.1609/aaai.v33i01.33012556>
- [11] Gangani Ariyaratne, Senuri De Silva, Sanuwani Dayarathna, Dulani Meedeniya, and Sampath Jayarathne. 2020. ADHD Identification using Convolutional Neural Network with Seed-based Approach for fMRI Data. 31–35. <https://doi.org/10.1145/3384544.3384552>
- [12] Aparna Balagopalan, Benjamin Eyre, Jessica Robin, Frank Rudzicz, and Jekaterina Novikova. 2021. Comparing Pre-trained and Feature-Based Models for Prediction of Alzheimer's Disease Based on Speech. *Frontiers in Aging Neuroscience* 13 (4 2021). <https://doi.org/10.3389/fnagi.2021.635945>
- [13] Azadeh Bashiri, Leila Shahmoradi, Hamid Beigy, Behrouz alizadeh savareh, Masoud Nosratabadi, Sharareh Rostam Niakan Kalhori, and Marjan Ghazisaeedi. 2018. Quantitative EEG features selection in the classification of attention and response control in the children and adolescents with attention deficit hyperactivity disorder. *Future Science OA* 4 (02 2018), FSO292. <https://doi.org/10.4155/fsoa-2017-0138>
- [14] Boller-F. Lopez O. L. Saxton J. McGonigle K. L. Becker, J. T. 1994. The natural history of Alzheimer's disease: description of study cohort and accuracy of diagnosis. *Archives of Neurology* 51(6) (1994), 585–594. <https://doi.org/10.1101/2021.03.24.21254263>
- [15] M. Bellani, A. Moretti, C. Perlini, and P. Brambilla. 2011. Language disturbances in ADHD. *Epidemiology and Psychiatric Sciences* 20 (12 2011), 311–315. Issue 4. <https://doi.org/10.1017/S2045796011000527>
- [16] Shristi Das Biswas, Rivu Chakraborty, and Ankita Pramanik. 2020. A Brief Survey on Various Prediction Models for Detection of ADHD from Brain-MRI Images. *Proceedings of the 21st International Conference on Distributed Computing and Networking*, 1–5. <https://doi.org/10.1145/3369740.3372775>
- [17] Lakshmi Boppana, Nikhat Shabnam, and Tadikonda Srivatsava. 2021. Deep Learning Approach for an early stage detection of Neurodevelopmental Disorders. 1–6. <https://doi.org/10.1109/R10-HTC53172.2021.9641691>
- [18] Barbro Bruce, Gunilla Thernlund, and Ulrika Nettelblad. 2006. ADHD and language impairment. *European Child Adolescent Psychiatry* 15 (2 2006), 52–60. Issue 1. <https://doi.org/10.1007/s00787-006-0508-9>
- [19] Carlton Chu Francois Chouinard Alan Evans András Jakab Budhachandra Singh Khundrakpam John David Lewis Qingyang Li Michael Milham Chaogan Yan Pierre Bellec Cameron Craddock, Yassine Benhajali. 2013. The Neuro Bureau Preprocessing Initiative: open sharing of preprocessed neuroimaging data and derivatives. *Neuroinformatics*. (2013).
- [20] Yang Chang, Cory Stevenson, I-Chun Chen, Dar-Shong Lin, and Li-Wei Ko. 2022. Neurological state changes indicative of ADHD in children learned via EEG-based LSTM networks. *Journal of Neural Engineering* 19 (01 2022). <https://doi.org/10.1088/1741-2552/ac4f07>
- [21] Nishant Chauhan and Byung-Jae Choi. 2020. DNN based Classification of ADHD fMRI Data using Functional Connectivity Coefficient. *International Journal of Fuzzy Logic and Intelligent Systems* 20 (12 2020), 255–260. <https://doi.org/10.5391/IJFIS.2020.20.4.255>
- [22] He Chen, Wenqing Chen, Yan Song, Li Sun, and Xiaoli Li. 2019. EEG characteristics of children with attention-deficit/hyperactivity disorder. *Neuroscience* 406 (5 2019), 444–456. <https://doi.org/10.1016/j.neuroscience.2019.03.048>
- [23] He Chen, Yan Song, and Xiaoli Li. 2019. A deep learning framework for identifying children with ADHD using an EEG-based brain network. *Neurocomputing* 356 (2019), 83–96. <https://doi.org/10.1016/j.neucom.2019.04.058>
- [24] He Chen, Yan Song, and Xiaoli Li. 2019. Use of deep learning to detect personalized spatial-frequency abnormalities in EEGs of children with ADHD. *Journal of Neural Engineering* 16, 6 (nov 2019), 066046. <https://doi.org/10.1088/1741-2552/ab3a0a>
- [25] Gulay Cicek and Aydin Akan. 2021. Deep Learning Approach Versus Traditional Machine Learning for ADHD Classification. 1–4. <https://doi.org/10.1109/TIPTEKNO53239.2021.9632940>
- [26] Natasha Clarke, Thomas R. Barrick, and Peter Garrard. 2021. A Comparison of Connected Speech Tasks for Detecting Early Alzheimer's Disease and Mild Cognitive Impairment Using Natural Language Processing and Machine Learning. *Frontiers in Computer Science* 3 (2021). <https://doi.org/10.3389/fcomp.2021.634360>
- [27] The Consortium. 2012. The ADHD-200 Consortium: a model to advance the translational potential of neuroimaging in clinical neuroscience. *Frontiers in Systems Neuroscience* 6 (2012). <https://doi.org/10.3389/fnsys.2012.00062>
- [28] Senuri De Silva, Sanuwani Dayarathna, Gangani Ariyaratne, Dulani Meedeniya, and Sampath Jayarathna. 2021. fMRI Feature Extraction Model for ADHD Classification Using Convolutional Neural Network. *International Journal of E-Health and Medical Communications* 12 (01 2021), 81–105. <https://doi.org/10.4018/IJEHMC.2021010106>
- [29] Victor Delvigne, Hazem Wannous, Thierry Dutoit, Laurence Ris, and Jean-Philippe Vandeboor. 2022. PhyDAA: Physiological Dataset Assessing Attention. *IEEE Transactions on Circuits and Systems for Video Technology* 32, 5 (2022), 2612–2623. <https://doi.org/10.1109/TCSVT.2021.3061719>
- [30] Qinglin Dong, Ning Qiang, Jinglei Lv, Xiang Li, Tianming Liu, and Quanzheng Li. 2020. Spatiotemporal Attention Autoencoder (STAAE) for ADHD Classification. , 508–517 pages. https://doi.org/10.1007/978-3-030-59728-3_50
- [31] Yao Dongren, Erkun Yang, Li Sun, Jing Sui, and Mingxia Liu. 2021. Integrating Multimodal MRIs for Adult ADHD Identification with Heterogeneous Graph Attention Convolutional Network. 157–167. https://doi.org/10.1007/978-3-030-87602-9_15
- [32] Kathleen C. Fraser, Kristina Lundholm Fors, and Dimitrios Kokkinakis. 2019. Multilingual word embeddings for the assessment of narrative speech in mild cognitive impairment. *Computer Speech Language* 53 (1 2019), 121–139. <https://doi.org/10.1016/j.csl.2018.07.005>
- [33] Julian Fritsch, Sebastian Wankerl, and Elmar Noth. 2019. Automatic Diagnosis of Alzheimer's Disease Using Neural Network Language Models. *ICASSP 2019 - 2019 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 5841–5845. <https://doi.org/10.1109/ICASSP.2019.8682690>
- [34] Ming-Shan Gao, Fu-Sheng Tsai, and Chi-Chun Lee. 2020. Learning a Phenotypic-Attribute Attentional Brain Connectivity Embedding for ADHD Classification using rs-fMRI. *2020 42nd Annual International Conference of the IEEE Engineering in Medicine Biology Society (EMBC)*, 5472–5475. <https://doi.org/10.1109/EMBC44109.2020.9175789>
- [35] Fabio Tamburini Gloria Gagliardi. 2021. Linguistic biomarkers for the detection of Mild Cognitive Impairment. *Lingue e linguaggio* (2021), 3–31. <https://doi.org/10.1418/101111>
- [36] Yuan Gong, Hasini Yatawatte, Christian Poellabauer, Sandra Schneider, and Susan Latham. 2018. Automatic Autism Spectrum Disorder Detection Using Everyday Vocalizations Captured by Smart Devices. *Proceedings of the 2018 ACM International Conference on Bioinformatics, Computational Biology, and Health Informatics*, 465–473. <https://doi.org/10.1145/3233547.3233574>
- [37] Gábor Gosztolya, Veronika Vincze, László Tóth, Magdolna Pákási, János Kálmán, and Ildikó Hoffmann. 2019. Identifying Mild Cognitive Impairment and mild Alzheimer's disease based on spontaneous speech using ASR and linguistic features. *Computer Speech Language* 53 (1 2019), 181–197. <https://doi.org/10.1016/j.csl.2018.07.007>
- [38] Ayşegül Güven, Miray Altınkaynak, Nazan Dolu, Meltem İzzetoğlu, Ferhat Pektaş, Sevgi Özmen, Esra Demirci, and Turgay Batbat. 2022. Correction To: Combining functional near-infrared spectroscopy and EEG measurements for the diagnosis of attention-deficit hyperactivity disorder. *Neural Computing and Applications* 34 (4 2022), 6581–6581. Issue 8. <https://doi.org/10.1007/s00521-022-07056-0>
- [39] Behshad Hosseinfard, Mohammad Hassan Moradi, and Reza Rostami. 2013. Classifying depression patients and normal subjects using machine learning techniques and nonlinear features from EEG signal. *Computer Methods and Programs in Biomedicine* 109 (3 2013), 339–345. Issue 3. <https://doi.org/10.1016/j.cmpb.2012.10.008>
- [40] Charles Lung-Cheng Huang, Shih-Feng Weng, and Chung-Han Ho. 2016. Gender ratios of administrative prevalence and incidence of attention-deficit/hyperactivity disorder (ADHD) across the lifespan: A nationwide population-based study in Taiwan. *Psychiatry Research* 244 (10 2016), 382–387. <https://doi.org/10.1016/j.psychres.2016.08.023>
- [41] Pratyaksh Jain, Karthik Ram Srinivas, and Abhishek Vichare. 2022. Depression and Suicide Analysis Using Machine Learning and NLP. *Journal of Physics: Conference Series* 2161 (1 2022), 012034. Issue 1. <https://doi.org/10.1088/1742-6596/2161/1/012034>
- [42] Ranjeet Jha, Aditya Nigam, Arnav Bhavsar, Gaurav Jaswal, Sudhir Pathak, and Rathish Kumar. 2020. Lightweight Graph Siamese Network with Triplet Loss for fMRI-based Classification of ADHD. <https://doi.org/10.1109/IJCNN48605.2020.9207667>
- [43] Ömer Kasım and Mustafa Tosun. 2021. Biometric Authentication from Photoc Stimulated EEG Records. *Applied Artificial Intelligence* (09 2021), 1–13. <https://doi.org/10.1080/08839514.2021.1981660>
- [44] Masahito Katsuki, Norio Narita, Yasuhiko Matsumori, Naoya Ishida, Ohmi Watanabe, Siqi Cai, and Teiji Tominaga. 2020. Preliminary development of a deep learning-based automated primary headache diagnosis model using Japanese natural language processing of medical questionnaire. *Surgical Neurology International* 11 (12 2020), 475. https://doi.org/10.25259/SNI_827_2020

- [45] Ali Khazaei, Ata Ebrahimzadeh, and Abbas Babajani-Feremi. 2016. Application of advanced machine learning methods on resting-state fMRI network for identification of mild cognitive impairment and Alzheimer's disease. *Brain Imaging and Behavior* 10 (9 2016), 799–817. Issue 3. <https://doi.org/10.1007/s11682-015-9448-7>
- [46] FLORENCE LEVY, DAVID A. HAY, KELLIE S. BENNETT, and MICHAEL MC-STEPHEN. 2005. Gender Differences in ADHD Subtype Comorbidity. *Journal of the American Academy of Child & Adolescent Psychiatry* 44 (4 2005), 368–376. Issue 4. <https://doi.org/10.1097/01.chi.0000153232.64968.c1>
- [47] Ming Li, Dengke Tang, Junlin Zeng, Tianyan Zhou, Huilin Zhu, Biyuan Chen, and Xiaobing Zou. 2019. An automated assessment framework for atypical prosody and stereotyped idiosyncratic phrases related to autism spectrum disorder. *Computer Speech & Language* 56 (7 2019), 80–94. <https://doi.org/10.1016/j.csl.2018.11.002>
- [48] Honglei Liu, Yan Xu, Zhiqiang Zhang, Ni Wang, Yanqun Huang, Yanjun Hu, Zhengnan Yang, Rui Jiang, and Hui Chen. 2020. A Natural Language Processing Pipeline of Chinese Free-Text Radiology Reports for Liver Cancer Diagnosis. *IEEE Access* 8 (2020), 159110–159119. <https://doi.org/10.1109/ACCESS.2020.3020138>
- [49] Ning Liu and Zhenming Yuan. 2021. Spontaneous Language Analysis in Alzheimer's Disease: Evaluation of Natural Language Processing Technique for Analyzing Lexical Performance. *Journal of Shanghai Jiaotong University (Science)* (11 2021). <https://doi.org/10.1007/s12204-021-2384-3>
- [50] Rui Liu, Zhi-an Huang, Min Jiang, and Kay Tan. 2020. Multi-LSTM Networks for Accurate Classification of Attention Deficit Hyperactivity Disorder from Resting-State fMRI Data. 1–6. <https://doi.org/10.1109/LAI50351.2020.9262176>
- [51] Shuaiqi Liu, Ling Zhao, Xu Wang, Qi Xin, Jie Zhao, David S. Guttery, and Yu-Dong Zhang. 2021. Deep Spatio-Temporal Representation and Ensemble Classification for Attention Deficit/Hyperactivity Disorder. *IEEE Transactions on Neural Systems and Rehabilitation Engineering* 29 (2021), 1–10. <https://doi.org/10.1109/TNSRE.2020.3019063>
- [52] Yu-Shuo Liu, Chin-Po Chen, Susan Shur-Fen Gau, and Chi-Chun Lee. 2018. Learning Lexical Coherence Representation Using LSTM Forget Gate for Children with Autism Spectrum Disorder During Story-Telling. *2018 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 6029–6033. <https://doi.org/10.1109/ICASSP.2018.8461560>
- [53] Ziming Liu, Lauren Proctor, Parker N. Collier, and Xiaopeng Zhao. 2021. Automatic Diagnosis and Prediction of Cognitive Decline Associated with Alzheimer's Dementia through Spontaneous Speech. In *2021 IEEE International Conference on Signal and Image Processing Applications (ICSIPA)*. 39–43. <https://doi.org/10.1109/ICSIPA52582.2021.9576784>
- [54] Hu Lu, Saixiong Liu, Hui Wei, Chao Chen, and Xia Geng. 2020. Deep multi-kernel auto-encoder network for clustering brain functional connectivity data. *Neural Networks* 135 (12 2020). <https://doi.org/10.1016/j.neunet.2020.12.005>
- [55] Tosun M. 2021. Effects of spectral features of EEG signals recorded with different channels and recording statuses on ADHD classification with deep learning. *Physical and engineering sciences in medicine* 2020, 693–702. <https://doi.org/10.1007/s13246-021-01018-x>
- [56] Abhay M S Aradhya and Andri Ashfahani. 2019. Deep Network Optimization for rs-fMRI Classification. In *2019 International Conference on Data Mining Workshops (ICDMW)*. 77–82. <https://doi.org/10.1109/ICDMW.2019.00022>
- [57] Pranav Mahajan and Veeky Baths. 2021. Acoustic and Language Based Deep Learning Approaches for Alzheimer's Dementia Detection From Spontaneous Speech. *Frontiers in Aging Neuroscience* 13 (2021). <https://doi.org/10.3389/fnagi.2021.623607>
- [58] Zhenyu Mao, Yi Su, Guangquan Xu, Xueping Wang, Yu Huang, Weihua Yue, Li Sun, and Naixue Xiong. 2019. Spatio-temporal deep learning method for ADHD fMRI classification. *Information Sciences* 499 (10 2019), 1–11. <https://doi.org/10.1016/j.ins.2019.05.043>
- [59] Jordi A. Matías-Guiu, Vanesa Pytel, Ana Cortés-Martínez, María Valles-Salgado, Teresa Rognoni, Teresa Moreno-Ramos, and Jorge Matías-Guiu. 2018. Conversion between Addenbrooke's Cognitive Examination III and Mini-Mental State Examination. *International Psychogeriatrics* 30 (8 2018), 1227–1233. Issue 8. <https://doi.org/10.1017/S104161021700268X>
- [60] Chris McNorgan, Cary Judson, Dakota Handzlik, and John G. Holden. 2020. Linking ADHD and Behavioral Assessment Through Identification of Shared Diagnostic Task-Based Functional Connections. *Frontiers in Physiology* 11 (12 2020). <https://doi.org/10.3389/fphys.2020.583005>
- [61] Lucas Miranda, Riya Paul, Benno Pütz, Nikolaos Koutsouleris, and Bertram Müller-Meyhok. 2021. Systematic Review of Functional MRI Applications for Psychiatric Disease Subtyping. *Frontiers in Psychiatry* 12 (10 2021), 665536. <https://doi.org/10.3389/fpsy.2021.665536>
- [62] Albert Moreira, Raul Alonso-Calvo, Alberto Muñoz, and José Crespo. 2018. Measuring Relevant Information in Health Social Network Conversations and Clinical Diagnosis Cases. *International Journal of Environmental Research and Public Health* 15 (12 2018), 2787. Issue 12. <https://doi.org/10.3390/ijerph15122787>
- [63] Muthuraman Muthuraman, Vera Moliadze, Lena Boecher, Julia Siemann, Christine M. Freitag, Sergiu Groppa, and Michael Siniatchkin. 2019. Multi-modal alterations of directed connectivity profiles in patients with attention-deficit/hyperactivity disorders. *Scientific Reports* 9 (12 2019), 20028. Issue 1. <https://doi.org/10.1038/s41598-019-56398-8>
- [64] Andreas Müller, Sarah Vetsch, Ilia Pershin, Gian Candrian, Gian-Marco Baschera, Juri D. Kropotov, Johannes Kasper, Hossam Abdel Rehim, and Dominique Eich. 2020. EEG/ERP-based biomarker/neuroalgorithms in adults with ADHD: Development, reliability, and application in clinical practice. *The World Journal of Biological Psychiatry* 21, 3 (2020), 172–182. <https://doi.org/10.1080/15622975.2019.1605198> PMID: 30990349.
- [65] Bahareh Nakisa, Naim Rastgoo, A. Rakotonirainy, Frederic Maire, and V. Chandran. 2018. Long Short Term Memory Hyperparameter Optimization for a Neural Network Based Emotion Recognition Framework. *IEEE Access* PP (09 2018), 1–1. <https://doi.org/10.1109/ACCESS.2018.2868361>
- [66] Ziad S. Nasreddine, Natalie A. Phillips, Valérie B. Bédard, Simon Charbonneau, Victor Whitehead, Isabelle Collin, Jeffrey L. Cummings, and Howard Chertkow. 2005. The Montreal Cognitive Assessment, MoCA: A Brief Screening Tool For Mild Cognitive Impairment. *Journal of the American Geriatrics Society* 53 (4 2005), 695–699. Issue 4. <https://doi.org/10.1111/j.1532-5415.2005.53221.x>
- [67] Shamila Nasreen, Morteza Rohanian, Julian Hough, and Matthew Purver. 2021. Alzheimer's Dementia Recognition From Spontaneous Speech Using Disfluency and Interactional Features. *Frontiers in Computer Science* 3 (6 2021). <https://doi.org/10.3389/fcomp.2021.640669>
- [68] Swati Padhee, Anurag Illendula, Megan Sadler, Valerie L. Shalin, Tanvi Banerjee, Krishnaprasad Thirunarayan, and William L. Romine. 2020. Predicting Early Indicators of Cognitive Decline from Verbal Utterances. In *2020 IEEE International Conference on Bioinformatics and Biomedicine (BIBM)*. 477–480. <https://doi.org/10.1109/BIBM49941.2020.9313106>
- [69] Raghavendra Pappagari, Jaemin Cho, Laureano Moro-Velázquez, and Najim Dehak. 2020. Using State of the Art Speaker Recognition and Natural Language Processing Technologies to Detect Alzheimer's Disease and Assess its Severity. 2177–2181. <https://doi.org/10.21437/Interspeech.2020-2587>
- [70] Ioana Petre. 2019. Explaining Gender Differences in the Diagnosis of ADHD in Children. , 183-197 pages. https://doi.org/10.1007/978-3-030-10677-5_12
- [71] Fariz Bramasta Putra, Alviansyah Arman Yusuf, Heri Yulianus, Yogi Putra Pratama, Dzakiyah Salma Humaira, Urriyatul Eriani, Dwi Kurnia Basuki, Sritrasta Sukaridhoto, and Rizqi Putri Nourma Budiarti. 2019. Identification of Symptoms Based on Natural Language Processing (NLP) for Disease Diagnosis Based on International Classification of Diseases and Related Health Problems (ICD-11). *2019 International Electronics Symposium (IES)*, 1–5. <https://doi.org/10.1109/ELECSYM.2019.8901644>
- [72] Ning Qiang, Qinglin Dong, Hongtao Liang, Bao Ge, Shu Zhang, Cheng Zhang, Jie Gao, and Yifei Sun. 2022. A novel ADHD classification method based on resting state temporal templates (RSTT) using spatiotemporal attention auto-encoder. *Neural Computing and Applications* 34 (5 2022), 7815–7833. Issue 10. <https://doi.org/10.1007/s00521-021-06868-w>
- [73] Khansa Rasheed, Adnan Qayyum, Junaid Qadir, Shobi Sivathamboo, Patrick Kwan, Levin Kuhlmann, Terence O'Brien, and Adeel Razi. 2021. Machine Learning for Predicting Epileptic Seizures Using EEG Signals: A Review. *IEEE Reviews in Biomedical Engineering* 14 (2021), 139–155. <https://doi.org/10.1109/RBME.2020.3008792>
- [74] Atif Riaz, Muhammad Asad, S M Masudur Rahman al Arif, E. Alonso, Danaí Dima, Philip Corr, and Greg Slabaugh. 2018. Deep fMRI: AN end-to-end deep network for classification of fMRI data. 1419–1422. <https://doi.org/10.1109/ISBI.2018.8363838>
- [75] Alireza Roshanzamir, Hamid Aghajan, and Mahdieh Soleymani Baghshah. 2021. Transformer-based deep neural network language models for Alzheimer's disease risk assessment from targeted speech. *BMC Medical Informatics and Decision Making* 21 (12 2021), 92. Issue 1. <https://doi.org/10.1186/s12911-021-01456-3>
- [76] S. de la Fuente D. Fromm B. MacWhinney S. Luz, F. Haider. 2021. Detecting cognitive decline using speech only: The ADReSSo Challenge. *medRxiv* 2021.03.24.21254263 (2021). <https://doi.org/10.1101/2021.03.24.21254263>
- [77] Sartaj Salman, Zhichao Lian, Marva Saleem, and Yuduo Zhang. 2020. Functional Connectivity Based Classification of ADHD Using Different Atlases. 62–66. <https://doi.org/10.1109/PIC50277.2020.9350749>
- [78] Larry J. Seidman, Eve M. Valera, and Nikos Makris. 2005. Structural Brain Imaging of Attention-Deficit/Hyperactivity Disorder. *Biological Psychiatry* 57 (6 2005), 1263–1272. Issue 11. <https://doi.org/10.1016/j.biopsych.2004.11.019>
- [79] Lizhen Shao, Donghui Zhang, Haipeng Du, and Dongmei Fu. 2019. Deep forest in ADHD data classification. *IEEE Access* PP (09 2019), 1–1. <https://doi.org/10.1109/ACCESS.2019.2941515>
- [80] Liu Shuaiqi, Ling Zhao, Jie Zhao, Bing Li, and Shui-Hua Wang. 2022. Attention deficit/hyperactivity disorder Classification based on deep spatio-temporal features of functional Magnetic Resonance Imaging. *Biomedical Signal Processing and Control* 71 (01 2022), 103239. <https://doi.org/10.1016/j.bspc.2021.103239>
- [81] Viktória Simon, Pál Czobor, Sára Bálint, Ágnes Mészáros, and István Bitter. 2009. Prevalence and correlates of adult attention-deficit hyperactivity disorder: meta-analysis. *British Journal of Psychiatry* 194 (3 2009), 204–211. Issue 3. <https://doi.org/10.1192/bjp.bp.107.048827>
- [82] Aradhana Soni, Benjamin Amrhein, Matthew Baucum, Eun Jin Paek, and Anahita Khojandi. 2021. Using Verb Fluency, Natural Language Processing, and Machine

- Learning to Detect Alzheimer's Disease. In *2021 43rd Annual International Conference of the IEEE Engineering in Medicine Biology Society (EMBC)*. 2282–2285. <https://doi.org/10.1109/EMBC46164.2021.9630371>
- [83] Yunkai Sun, Lei Zhao, Zhihui Lan, Xi-Ze Jia, and Shao-Wei Xue. 2020. <p>Differentiating Boys with ADHD from Those with Typical Development Based on Whole-Brain Functional Connections Using a Machine Learning Approach</p>. *Neuropsychiatric Disease and Treatment* Volume 16 (3 2020), 691–702. <https://doi.org/10.2147/NDT.S239013>
- [84] Amirali Vahid, Annet Bluschke, Veit Roessner, Sebastian Stober, and Christian Beste. 2019. Deep Learning Based on Event-Related EEG Differentiates Children with ADHD from Healthy Controls. *Journal of Clinical Medicine* 8 (7 2019), 1055. Issue 7. <https://doi.org/10.3390/jcm8071055>
- [85] Tianyi Wang and Sei ichiro Kamata. 2019. Classification of Structural MRI Images in Adhd Using 3D Fractal Dimension Complexity Map. *2019 IEEE International Conference on Image Processing (ICIP)*, 215–219. <https://doi.org/10.1109/ICIP.2019.8802930>
- [86] Xun-Heng Wang, Jie Xu, and Lihua Li. 2020. Estimating individual scores of inattention and impulsivity based on dynamic features of intrinsic connectivity network. *Neuroscience Letters* 724 (02 2020), 134874. <https://doi.org/10.1016/j.neulet.2020.134874>
- [87] Zijian Wang, Yaqin Zhu, Haibo Shi, Yanting Zhang, and Cairong Yan. 2021. A 3D multiscale view convolutional neural network with attention for mental disease diagnosis on MRI images. *Mathematical Biosciences and Engineering* 18 (08 2021), 6978–3994. <https://doi.org/10.3934/mbe.2021347>
- [88] Dongren Yao, Xiaojie Guo, Qihua Zhao, Lu Liu, Qingjiu Cao, Yufeng Wang, Vince D Calhoun, Li Sun, and Jing Sui. 2018. Discriminating ADHD From Healthy Controls Using a Novel Feature Selection Method Based on Relative Importance and Ensemble Learning. *2018 40th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, 4632–4635. <https://doi.org/10.1109/EMBC.2018.8513155>
- [89] Zi Ye, Shoukang Hu, Jinchao Li, Xurong Xie, Mengzhe Geng, Jianwei Yu, Junhao Xu, Boyang Xue, Shansong Liu, Xunying Liu, and Helen Meng. 2021. Development of the Cuhk Elderly Speech Recognition System for Neurocognitive Disorder Detection Using the Dementiabank Corpus. *ICASSP 2021 - 2021 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 6433–6437. <https://doi.org/10.1109/ICASSP39728.2021.9413634>
- [90] Yangsong Zhang, Huan Cai, Li Nie, Peng Xu, Sirui Zhao, and Cuntai Guan. 2021. An end-to-end 3D convolutional neural network for decoding attentive mental state. *Neural Networks* 144 (08 2021). <https://doi.org/10.1016/j.neunet.2021.08.019>
- [91] Zirui Zhang and Xiaoli Li. 2019. Use transfer learning to promote identification ADHD children with EEG recordings. *2019 Chinese Automation Congress (CAC)*, 2809–2813. <https://doi.org/10.1109/CAC48633.2019.8997426>
- [92] Yanli Zhang-James, Emily Helminen, Jinru Liu, Barbara Franke, Martine Hoogman, and Stephen Faraone. 2021. Evidence for similar structural brain anomalies in youth and adult attention-deficit/hyperactivity disorder: a machine learning analysis. *Translational Psychiatry* 11 (02 2021). <https://doi.org/10.1038/s41398-021-01201-4>
- [93] Kanhao Zhao, Boris Duka, Hua Xie, Desmond Oathes, Vince Calhoun, and Yu Zhang. 2022. A dynamic graph convolutional neural network framework reveals new insights into connectome dysfunctions in ADHD. *NeuroImage* 246 (02 2022), 118774. <https://doi.org/10.1016/j.neuroimage.2021.118774>
- [94] Qiu-Yue Zhong, Elizabeth W. Karlson, Bizu Gelaye, Sean Finan, Paul Avillach, Jordan W. Smoller, Tianxi Cai, and Michelle A. Williams. 2018. Screening pregnant women for suicidal behavior in electronic medical records: diagnostic codes vs. clinical notes processed by natural language processing. *BMC Medical Informatics and Decision Making* 18 (12 2018), 30. Issue 1. <https://doi.org/10.1186/s12911-018-0617-7>
- [95] Liang Zou, Jiannan Zheng, and Martin McKeown. 2017. Deep learning based automatic diagnoses of attention deficit hyperactive disorder. 962–966. <https://doi.org/10.1109/GlobalSIP.2017.8309103>

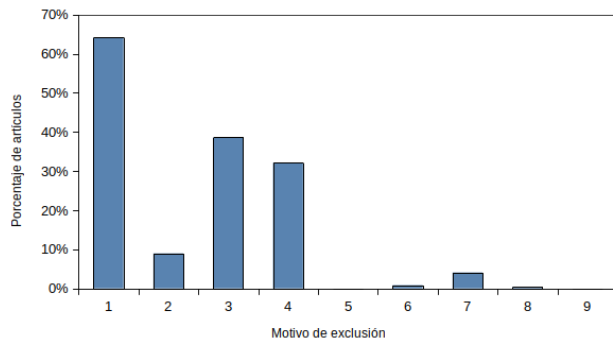


Figura 6: Distribución de motivos de exclusión. Se muestra el porcentaje de artículos que cumplen con cada motivo de exclusión considerado. Se consideró que varios motivos de exclusión pueden aplicar al mismo artículo, por lo que la suma de porcentajes supera el 100%. Los principales motivos de exclusión fueron: 1) que los artículos no estudiaran la enfermedad adecuada, 2) que no utilizaran modelos de aprendizaje profundo y 3) que su objetivo no fuera la predicción de la enfermedad.

7 APPENDICES

7.1 Apéndice I: Estudio Bibliográfico

La revisión bibliográfica llevada a cabo en este trabajo se organizó en las siguientes etapas:

- 1) Búsqueda exploratoria en Google Scholar en base a la cuál se identificaron las palabras clave asociadas a los artículos de interés.
- 2) Obtención de un listado inicial de 356 artículos a través de búsquedas en las siguientes bases de datos bibliográficas: 1) **Web of Science** (225 artículos), 2) **Scopus** (263 artículos) y 3) **IEEE Xplore** (41 artículos). Las consultas utilizadas en cada base de datos se describen en el Apéndice 7.2
- 3) Filtrado de 110 artículos de interés (sólo 106 fueron finalmente accesibles) mediante lectura y filtrado de *abstracts*. Los criterios de exclusión se detallan en el Apéndice 7.3 y la Figura 6 muestra el porcentaje de artículos que cumplía los distintos criterios.
- 4) Lectura completa de los artículos finales y elaboración de resúmenes y una tabla comparativa (Tabla 2).

7.2 Apéndice II: Consultas en bases de datos bibliográficas

7.3 Apéndice III: Criterios de exclusión

Los criterios de exclusión considerados fueron:

- 1) **Enfermedad:** utiliza EEG/MRI como fuente de datos pero estudia una enfermedad que no es TDAH.
- 2) **PLN:** utiliza PLN como fuente de datos pero no diagnóstica en función de características lingüísticas del paciente.
- 3) **Modelo:** utiliza EEG/MRI como fuente de datos pero el clasificador aplicado no es de aprendizaje profundo, ya que el análisis exploratorio reveló que estos modelos suponen el estado del arte actual.

```
{
{
((TI="ADHD" OR TS="ADHD") OR
(TI="attention deficit" OR TS="attention deficit") OR
(TI="hyperactiv" OR TS="hyperactiv"))
AND
((TI="magnetic resonance" OR TS="magnetic resonance") OR
(TI="MRI" OR TS="MRI") OR
(TI="fMRI" OR TS="fMRI") OR
(TI="magnetoencephalo" OR TS="magnetoencephalo") OR
(TI="MEG" OR TS="MEG") OR
(TI="electroencephal" OR TS="electroencephal") OR
(TI="EEG" OR TS="EEG"))
AND
((TI="ML" OR TS="ML") OR
(TI="machine learn" OR TS="machine learn") OR
(TI="deep learn" OR TS="deep learn") OR
(TI="deep neural net" OR TS="deep neural net") OR
(TI="DNN" OR TS="DNN") OR
(TI="CNN" OR TS="CNN") OR
((TI="convolut" OR TS="convolut") AND
(TI="neural net" OR TS="neural net")) OR
(TI="LSTM" OR TS="LSTM") OR
(TI="RNN" OR TS="RNN") OR
(TI="recurrent neural net" OR TS="recurrent neural net"))
) OR (
((TI="diagnos" OR TS="diagnos") AND
((TI="ADHD" OR TS="ADHD") OR
(TI="attention deficit" OR TS="attention deficit") OR
(TI="hyperactiv" OR TS="hyperactiv") OR
(TI="alzheimer" OR TS="alzheimer") OR
(TI="autis" OR TS="autis") OR
(TI="ADS" OR TS="ADS"))))
AND
((TI="NLP" OR TS="NLP") OR
(TI="natural language process" OR TS="natural language process") OR
(TI="text analy" OR TS="text analy") OR
(TI="text min" OR TS="text min"))
NOT
((TI="clinical note" OR TS="clinical note") OR
(TI="health record" OR TS="health record") OR
(TI="web" OR TS="web") OR
(TI="social medi" OR TS="social medi") OR
(TI="social net" OR TS="social net") OR
(TI="gen" OR TS="gen") OR
(TI="pheno" OR TS="pheno"))
)
) AND (PY=("2022" OR "2021" OR "2020" OR "2019" OR "2018"))
}
```

Figura 7: Consulta Wos.

```
{
{
TITLE-ABS-KEY ( "ADHD" OR "attention deficit" OR "hyperactiv" ) AND
TITLE-ABS-KEY ( "magnetic resonance" OR "magnetic resonance" OR "MRI" OR
"fMRI" OR "fMRI" OR "FMI" OR "magnetoencephalo" OR "MEG" OR "electroencephal" OR
"EEG" ) AND
TITLE-ABS-KEY ( "machine learn" OR "deep learn" OR "deep neural net" OR
"deep neural net" OR "DNN" OR "CNN" OR "convolut" OR "LSTM" OR "RNN" OR
"recurrent neural net" )
) OR (
TITLE-ABS-KEY ( "diagnos" ) AND
TITLE-ABS-KEY ( "ADHD" OR "attention deficit" OR "hyperactiv" OR
"alzheimer" OR "autis" OR "ADS" ) AND
TITLE-ABS-KEY ( "NLP" OR "natural language process" OR "text analy" OR
"text min" ) AND NOT
TITLE-ABS-KEY ( "clinical not" OR "health record" OR "web" OR "social
medi" OR "social net" OR "gen" OR "pheno" )
)
) AND ( PUBYEAR > 2017 AND
(LIMIT-TO ( DOCTYPE , "ar" ) OR
LIMIT-TO ( DOCTYPE , "cp" ) OR
LIMIT-TO ( DOCTYPE , "re" ) ) )
}
```

Figura 8: Consulta Scopus.

- 4) **Objetivo:** utiliza EEG/MRI/PLN como fuente de datos para el diagnóstico, en lugar de solo mencionarlas o comentar que tienen una influencia sobre el mismo.
- 5) **Conjunto de datos:** se aceptan todos los conjuntos de datos, pero en caso de que sea necesario reducir más el número de aceptados se priorizarán artículos con conjuntos más extensos o diversos.
- 6) **Seminarios:** el artículo corresponde a un seminario.
- 7) **Revisiones:** revisa la bibliografía de los temas correctos (véase 1., 2. y 4.), pero no aporta nada respecto a los demás artículos ya aceptados (u otras revisiones).

```

(
  ("Document Title":ADHD OR "Abstract":ADHD) OR
  ("Document Title":attention deficit OR "Abstract":attention deficit) OR
  ("Document Title":hyperactivity OR "Abstract":hyperactivity)
) AND (
  ("Document Title":magnetic resonance OR "Abstract":magnetic resonance) OR
  ("Document Title":MRI OR "Abstract":MRI) OR
  ("Document Title":fMRI OR "Abstract":fMRI) OR
  ("Document Title":magnetoencephalography OR
  "Abstract":magnetoencephalography) OR
  ("Document Title":MEG OR "Abstract":MEG) OR
  ("Document Title":electroencephalography OR "Abstract":electroencephalography)
OR
  ("Document Title":EEG OR "Abstract":EEG)
) AND (
  ("Document Title":ML OR "Abstract":ML) OR
  ("Document Title":machine learning OR "Abstract":machine learning) OR
  ("Document Title":DL OR "Abstract":DL) OR
  ("Document Title":deep learning OR "Abstract":deep learning) OR
  ("Document Title":neural network OR "Abstract":neural network) OR
  ("Document Title":NN OR "Abstract":NN) OR
  ("Document Title":deep neural network OR "Abstract":deep neural network) OR
  ("Document Title":DNN OR "Abstract":DNN) OR
  ("Document Title":CNN OR "Abstract":CNN) OR
  ( ("Document Title":convolution OR "Abstract":convolution) AND
  ("Document Title":neural network OR "Abstract":neural network) ) OR
  ("Document Title":LSTM OR "Abstract":LSTM) OR
  ("Document Title":RNN OR "Abstract":RNN) OR
  ("Document Title":recurrent neural network OR "Abstract":recurrent neural
  network)
)

(
  ("Document Title":diagnos* OR "Abstract":diagnos*) AND
  ( ("Document Title":ADHD OR "Abstract":ADHD) OR
  ("Document Title":attention deficit OR "Abstract":attention deficit) OR
  ("Document Title":hyperactivity OR "Abstract":hyperactivity) OR
  ("Document Title":alzheimer OR "Abstract":alzheimer) OR
  ("Document Title":autism OR "Abstract":autism) OR
  ("Document Title":ADS OR "Abstract":ADS) )
) AND (
  ("Document Title":NLP OR "Abstract":NLP) OR
  ("Document Title":natural language processing OR "Abstract":natural language
  processing) OR
  ("Document Title":text analysis OR "Abstract":text analysis) OR
  ("Document Title":text mining OR "Abstract":text mining)
) NOT (
  ("Document Title":clinical note* OR "Abstract":clinical note*) OR
  ("Document Title":health record* OR "Abstract":health record*) OR
  ("Document Title":web OR "Abstract":web) OR
  ("Document Title":social media OR "Abstract":social media) OR
  ("Document Title":social network OR "Abstract":social network) OR
  ("Document Title":gen OR "Abstract":gen) OR
  ("Document Title":phenotype OR "Abstract":phenotype)
)

```

Figura 9: Consulta IEEE. Esta consulta se dividió en dos partes y se obtuvo la unión de los resultados aportados por cada una.

(8) **No artículo:** no corresponde a un artículo de investigación ni revisión (editoriales, *open letters*, *viewpoints*, ...).

(9) **Otros motivos.**

Pese a que un criterio de exclusión en el filtrado de la bibliografía fue no utilizar modelos que no fueran de aprendizaje profundo, el volumen de trabajos con modelos clásicos era tan alto que no todos se pudieron identificar con la lectura de *abstracts*. El principal problema de estos clasificadores es que las redes profundas obtienen resultados mejores consistentemente, especialmente en el campo de análisis de imagen.

7.4 Apéndice IV: Resumen bibliográfico

Siglas utilizadas en la Tabla 2:

- ESQ: esquizofrenia
- PTSD: síndrome de estrés postraumático
- NEU: cualquier enfermedad neurológica
- C/R: clasificación (C) o regresión (R)

Tabla 2: Resumen del estado del arte en predicción de TDAH mediante neuroimagen y de otras enfermedades mediante PLN.

Título	Año	Enfermedad	C/R	Fuente de datos	Conjunto de datos	Actividad	Tamaño	Preprocesamiento	Modelo	Precisión
Neuroimaging in attention-deficit/hyperactivity disorder	2020	TDAH	-	-	-	-	-	-	-	-
A deep learning framework for identifying children with adhd using an eeg-based brain network	2019	TDAH	C	EEG	Propio	Reposo	107	S	CNN	94,67%
A review of eeg signal analysis for diagnosis of neurological disorders using machine learning	2021	TDAH	C	EEG	-	-	-	S/N	SVM CNN	84,44%
An approach to measure and improve the cognitive capability of adhd affected children through eeg signals	2018	TDAH	C	EEG	Propio	Otros	-	S	-	-
An end-to-end 3d convolutional neural network for decoding attentive mental state	2021	TDAH	C	EEG	Propio	AOT	24	S	CNN	77,00%
Biometric authentication from photic stimulated eeg records	2021	TDAH	C	EEG	Propio	Otros	16	S	CNN	97,00%
Combining functional near-infrared spectroscopy and eeg measurements for the diagnosis of attention-deficit hyperactivity disorder	2019	TDAH	C	EEG	Propio	AOT	44	S	SVM NN	95,60%
Computer aided diagnosis system using deep convolutional neural networks for adhd subtypes	2021	TDAH	C	EEG	Propio	Reposo	39	S	CNN	97,20%
Deep learning based on event-related eeg differentiates children with adhd from healthy controls	2019	TDAH	C	EEG	Propio	AOT	144	N	CNN	83,00%
Deep learning convolutional neural networks discriminate adult adhd from healthy individuals on the basis of event-related spectral eeg	2020	TDAH	C	EEG	Propio	Otros	40	S	CNN	88,00%
Diagnosis of attention deficit hyperactivity disorder with combined time and frequency features	2020	TDAH	C	EEG	Propio	AOT	46	S	NN	91,00%
Differences in performance of asd and adhd subjects facing cognitive loads in an innovative reasoning experiment	2021	TEA TDAH	-	EEG	Propio	Otros	63	S	-	-

Tabla 2: Resumen del estado del arte en predicción de TDAH mediante neuroimagen y de otras enfermedades mediante PLN.

Título	Año	Enfermedad	C/R	Fuente de datos	Conjunto de datos	Actividad	Tamaño	Preprocesamiento	Modelo	Precisión
EEG characteristics of children with attention-deficit/hyperactivity disorder	2019	TDAH	C	EEG	Propio	Reposo	108	S	SVM	-
EEG spectral power, but not theta/beta ratio, is a neuromarker for adult adhd	2019	TDAH	C	EEG	Propio	Reposo	134	S	Otros	56,00%
EEG/ERP - biomarker/neuroalgorithms in adults with adhd: development, reliability and application in clinical practice	2019	TDAH	C	EEG	Propio	Reposo	328	S	-	83,00%
Effects of spectral features of eeg signals recorded with different channels and recording statuses on adhd classification with deep learning	2021	TDAH	C	EEG	Propio	Reposo	17	S	RNN	92,15%
Focus: detecting adhd patients by an eeg-based serious game	2018	TDAH	C	EEG	Propio	AOT	4	S	SVM	98,62%
LSTM hyperparameter optimization for an eeg-based efficient emotion classification in bci	2018	-	C	EEG	Propio	Otros	20	S	RNN	77,68%
Methodology proposal of adhd classification of children based on cross recurrence plots	2021	TDAH	C	EEG	Propio	Reposo	121	S	-	94,93%
Multimodal alterations of directed connectivity profiles in patients with attention-deficit/hyperactivity disorders	2019	TDAH	C	EEG	Propio	Reposo	22	S	-	98,00%
Neurological state changes indicative of adhd in children learned via eeg-based lstm networks.	2022	TDAH	C	EEG	Propio	Reposo	60	S	RNN	90,00%
Objective adhd diagnosis using convolutional neural networks over daily-life activity records	2020	TDAH	C	EEG	Propio	-	148	S	CNN	98,77%
PHYDAA: physiological dataset assessing attention	2021	TDAH	C	EEG	Propio	Otros	32	S	CNN RNN	85,80%

Tabla 2: Resumen del estado del arte en predicción de TDAH mediante neuroimagen y de otras enfermedades mediante PLN.

Título	Año	Enfermedad	C/R	Fuente de datos	Conjunto de datos	Actividad	Tamaño	Preprocesamiento	Modelo	Precisión
Quantitative eeg features selection in the classification of attention and response control in the children and adolescents with attention deficit hyperactivity disorder	2018	TDAH	C	EEG	Propio	Reposo	95	S	NN	100,00%
Use of deep learning to detect personalized spatial-frequency abnormalities in eegs of children with adhd	2019	TDAH	C	EEG	Propio	Reposo	107	S	CNN	90,00%
Use transfer learning to promote identification adhd children with eeg recordings	2019	TDAH	C	EEG	Propio	-	108	S	CNN	97,83%
A 3d multiscale view convolutional neural network with attention for mental disease diagnosis on mri images	2021	TDAH ESQ	C	fMRI	ADHD-200	Reposo	800	N	CNN	78,80%
A brief survey on various prediction models for detection of adhd from brain-mri images	2020	TDAH	C	fMRI	ADHD-200	Reposo	937	S	SVM CNN RNN CNN	- 72,00%
A dynamic graph convolutional neural network framework reveals new insights into connectome dysfunctions in adhd	2022	TDAH	C	fMRI	ADHD-200	Reposo	635	S		
A general prediction model for the detection of adhd and autism using structural and functional mri	2018	TEA TDAH	C	fMRI	ADHD-200 ABIDE	Reposo	1.045	S	SVM	-
A network-based approach to study of adhd using tensor decomposition of resting state fmri data	2020	TDAH	C	fMRI	ADHD-200	Reposo	259	S	SVM	79,60%
A novel adhd classification method based on resting state temporal templates (rstt) using spatiotemporal attention auto-encoder	2022	TDAH	C	fMRI	ADHD-200	Reposo	937	S	AE+NN	90,00%
ADHD classification using auto-encoding neural network and binary hypothesis testing	2022	TDAH	C	fMRI	ADHD-200	Reposo		S	SVM	99,60%
ADHD identification using convolutional neural network with seed-based approach for fmri data	2020	TDAH	C	fMRI	ADHD-200	Reposo	627	S	CNN	85,21%

Tabla 2: Resumen del estado del arte en predicción de TDAH mediante neuroimagen y de otras enfermedades mediante PLN.

Título	Año	Enfermedad	C/R	Fuente de datos	Conjunto de datos	Actividad	Tamaño	Preprocesamiento	Modelo	Precisión
Application of artificial intelligence in the mri classification task of human brain neurological and psychiatric diseases: a scoping review	2021	NEU	-	fMRI	-	-	40	-	-	-
Attention deficit/hyperactivity disorder classification based on deep spatio-temporal features of functional magnetic resonance imaging	2022	TDAH	C	fMRI	ADHD-200	Reposo	937	S	AE+NN	72,44%
Brain mr image classification for adhd diagnosis using deep neural networks	2020	TDAH	C	fMRI	ADHD-200	Reposo	937	S	CNN	75,00%
Classification of structural mri images in adhd using 3d fractal dimension complexity map	2019	TDAH	C	fMRI	ADHD-200	Reposo	937	S	CNN	60,90%
Deep fmri: an end-to-end deep network for classification of fmri data	2018	TDAH	C	fMRI	ADHD-200	Reposo	359	S	NN	73,10%
Deep forest in adhd data classification	2019	TDAH	C	fMRI	ADHD-200	Reposo	540	S	RF	82,73%
Deep learning approach for an early stage detection of neurodevelopmental disorders	2021	TEA TDAH ESQ	C	fMRI	ABIDE ADHD-200 COBRE UCLA WUSTL	Reposo	1.200	S	CNN	71,16%
Deep learning approach versus traditional machine learning for adhd classification	2021	TDAH	C	fMRI	ADHD-200	Reposo	937	S	SVM CNN	100,00%
Deep learning based automatic diagnoses of attention deficit hyperactive disorder	2017	TDAH	C	fMRI	ADHD-200	Reposo	171	S	CNN	66,00%
Deep multi-kernel auto-encoder network for clustering brain functional connectivity data	2020	EA TDAH PSTD	A	fMRI	ABIDE	Reposo	1.088	S	NN	63,00%
Deep network optimization for rs-fmri classification	2019	TDAH	C	fMRI	ADHD-200	Reposo	937	N*	CNN	80,39%
Deep neural generative adversarial model based on vae + gan for disorder diagnosis	2020	TEA TDAH	DA	fMRI	ADHD-200 ABIDE	Reposo	941	S	AE GAN	-

Tabla 2: Resumen del estado del arte en predicción de TDAH mediante neuroimagen y de otras enfermedades mediante PLN.

Título	Año	Enfermedad	C/R	Fuente de datos	Conjunto de datos	Actividad	Tamaño	Preprocesamiento	Modelo	Precisión
Deep spatio-temporal representation and ensemble classification for attention deficit/hyperactivity disorder	2021	TDAH	C	fMRI	ADHD-200	Reposo	937	S	CNN	82,00%
Deep transformation method for discriminant analysis of multi-channel resting state fmri	2019	TDAH	C	fMRI	ADHD-200	Reposo	465	S	CNN	70,36%
Deep variational autoencoder for mapping functional brain networks	2021	TDAH	C	fMRI	ADHD-200	Reposo	937	S	AE+NN	75,00%
DeepfMRI: end-to-end deep learning for functional connectivity and classification of adhd using fmri	2022	TDAH	C	fMRI	ADHD-200	Reposo	937	S	CNN	91,60%
Differentiating boys with adhd from those with typical development based on whole-brain functional connections using a machine learning approach	2020	TDAH	C	fMRI	Propio	Reposo	68	S	-	87,00%
Dilated 3d convolutional neural networks for brain mri data classification	2018	TDAH ESQ	C	fMRI	ADHD-200	Reposo	937	S	CNN	76,00%
Discriminating adhd from healthy controls using a novel feature selection method based on relative importance and ensemble learning	2018	TDAH	C	fMRI	Propio	-	107	S	RF	86,00%
DNN based classification of adhd fmri data using functional connectivity coefficient	2020	TDAH	C	fMRI	ADHD-200	Reposo	937	S	NN	95,00%
Estimating gender and age from brain structural mri of children and adolescents: a 3d convolutional neural network multitask learning model	2021	TDAH	C R	fMRI	ADHD-200	Reposo	937	S	CNN	-
Estimating individual scores of inattention and impulsivity based on dynamic features of intrinsic connectivity network	2020	TDAH	R	fMRI	ADHD-200	Reposo	210	S	-	-
Evidence for similar structural brain anomalies in youth and adult attention-deficit/hyperactivity disorder: a machine learning analysis	2021	TDAH	C	fMRI	ADHD-200	Reposo	776	S	NN	68,00%

Tabla 2: Resumen del estado del arte en predicción de TDAH mediante neuroimagen y de otras enfermedades mediante PLN.

Título	Año	Enfermedad	C/R	Fuente de datos	Conjunto de datos	Actividad	Tamaño	Preprocesamiento	Modelo	Precisión
fMRI feature extraction model for adhd classification using convolutional neural network	2021	TDAH	C	fMRI	ADHD-200	Reposo	40	S	CNN	85,36%
Functional connectivity based classification of adhd using different atlases	2020	TDAH	C	fMRI	ADHD-200	Reposo	243	S	AE+NN	99,00%
Fusion of fmri and non-imaging data for adhd classification	2017	TDAH	C	fMRI	ADHD-200	Reposo	937	S	SVM	69,00%
Graph deep learning-based attention-deficit/hyperactivity disorder diagnosis for brain network	2021	TDAH	C	fMRI	Propio	-	456	S	CNN	72,10%
HLGsNET: hierarchical and lightweight graph siamese network with triplet loss for fmri-based classification of adhd	2020	TDAH	C	fMRI	ADHD-200	Reposo	937	S	CNN	91,60%
Integrating multimodal mris for adult adhd identification with heterogeneous graph attention convolutional network	2021	TDAH	C	fMRI	Propio	Reposo	187	S	NN	76,30%
Investigating the correspondence of clinical diagnostic grouping with underlying neurobiological and phenotypic clusters using unsupervised machine learning	2018	TDAH	C	fMRI	ADHD-200	Reposo	937	S	-	69,34%
Latent brain state dynamics distinguish behavioral variability, impaired decision-making, and inattention	2021	TDAH	-	fMRI	-	-	-	-	-	-
Learning a phenotypic-attribute attentional brain connectivity embedding for adhd classification using rs-fmri	2020	TDAH	C	fMRI	ADHD-200	Reposo	937	S	AE+NN	86,22%
Linking adhd and behavioral assessment through identification of shared diagnostic task-based functional connections	2020	TDAH	C	fMRI	MTA-168	Otros	80	S	NN	91,00%

Tabla 2: Resumen del estado del arte en predicción de TDAH mediante neuroimagen y de otras enfermedades mediante PLN.

Título	Año	Enfermedad	C/R	Fuente de datos	Conjunto de datos	Actividad	Tamaño	Preprocesamiento	Modelo	Precisión
Machine learning methods for diagnosing autism spectrum disorder and attention- deficit/hyperactivity disorder using functional and structural mri: a survey	2021	TEA TDAH	C	fMRI	-	-	-	-	-	-
Multi-LSTM networks for accurate classification of attention deficit hyperactivity disorder from resting-state fmri data	2020	TDAH	C	fMRI	ADHD-200	Reposo	937	S	RNN	73,00%
Promises and pitfalls of deep neural networks in neuroimaging-based psychiatric research	2021	EA TEA TDAH ESQ	C	fMRI	Propio	CT	-	S	SVM CNN RNN	90,00%
Separated channel attention convolutional neural network (sc-cnn-attention) to identify adhd in multi-site rs-fmri dataset	2020	TDAH	C	fMRI	ADHD-200	Reposo	422	S	CNN RNN	68,60%
Spatio-temporal deep learning method for adhd fmri classification	2019	TDAH	C	fMRI	ADHD-200	Reposo	941	S	CNN RNN	71,00%
Spatiotemporal attention autoencoder (staae) for adhd classification	2020	TDAH	C	fMRI	ADHD-200	Reposo	937	S	AE+NN	70,00%
Systematic review of functional mri applications for psychiatric disease subtyping	2021	NEU	C	fMRI	-	-	-	S	-	-
A survey of attention deficit hyperactivity disorder identification using psychophysiological data	2019	TDAH	C	fMRI EEG	-	-	-	S	-	-
A comparison of connected speech tasks for detecting early alzheimer's disease and mild cognitive impairment using natural language processing and machine learning	2021	EA	C	PLN	Propio	CT	50	S	-	78,00%
A review of automated speech and language features for assessment of cognitive and thought disorders	2019	EA	-	PLN	-	-	-	-	-	-
Acoustic and language based deep learning approaches for alzheimer's dementia detection from spontaneous speech	2021	EA	C	PLN	ADReSSo	E	156	S	RNN CNN+RNN NN	73,00%

Tabla 2: Resumen del estado del arte en predicción de TDAH mediante neuroimagen y de otras enfermedades mediante PLN.

Título	Año	Enfermedad	C/R	Fuente de datos	Conjunto de datos	Actividad	Tamaño	Preprocesamiento	Modelo	Precisión
Alzheimer's dementia recognition from spontaneous speech using disfluency and interactional features	2021	EA	C	PLN	CCC	E	30	S	SVM NN	83,00%
An automated assessment framework for atypical prosody and stereotyped idiosyncratic phrases related to autism spectrum disorder	2018	TEA	C	PLN	Propio	E	189	S	CNN RNN CNN+RNN	88,10%
Automatic autism spectrum disorder detection using everyday vocalizations captured by smart devices	2018	TEA	C	PLN	Propio	Otros	35	S	SVM	87,00%
Automatic diagnosis and prediction of cognitive decline associated with alzheimer's dementia through spontaneous speech	2021	EA	C	PLN	ADReSSo	CT	239	S	BERT	97,00%
Automatic diagnosis of alzheimer's disease using neural network language models	2019	EA	C	PLN	DementiaBankCT		292	N	RNN	85,60%
Comparing pre-trained and feature-based models for prediction of alzheimer's disease based on speech	2021	EA	C	PLN	ADReSSo	CT	156	S	SVMRF NB NN BERT	83,32%
Computerized neuropsychological assessment in mild cognitive impairment based on natural language processing-oriented feature extraction	2017	EA	R	PLN	Propio	D	3	S	-	-
Computing the structure of language for neuropsychiatric evaluation	2017	EA	-	PLN	Propio	-	-	N	-	-
Correlating natural language processing and automated speech analysis with clinician assessment to quantify speech-language changes in mild cognitive impairment and alzheimer's dementia	2021	EA	C	PLN	DementiaBankCT		30	S	-	86,00%
Detecting alzheimer's disease by exploiting linguistic information from nepali transcript	2022	EA	C	PLN	Propio	CT	266	S	CNN RNN CNN+RNN	96,80%
Detection of mild cognitive impairment through natural language and touchscreen typing processing.	2020	EA	C	PLN	Propio	Otros	23	S	CNN	78,00%

Tabla 2: Resumen del estado del arte en predicción de TDAH mediante neuroimagen y de otras enfermedades mediante PLN.

Título	Año	Enfermedad	C/R	Fuente de datos	Conjunto de datos	Actividad	Tamaño	Preprocesamiento	Modelo	Precisión
Development of the cuhk elderly speech recognition system for neurocognitive disorder detection using the dementiaBank corpus	2021	EA	C	PLN	DementiaBankCT		292	S	BERT	88,00%
How to do things with (thousands of) words: computational approaches to discourse analysis in alzheimer's disease	2020	EA	-	PLN	-	-	-	-	-	-
Identifying mild cognitive impairment and mild alzheimer's disease based on spontaneous speech using asr and linguistic features	2019	EA	C	PLN	Propio	E	75	S	SVM NN	86,00%
Is cognitive impairment associated with reduced syntactic complexity in writing? evidence from automated text analysis	2020	EA	R	PLN	Propio	CT	138	S	-	-
Learning converse-level multimodal embedding to assess social deficit severity for autism spectrum disorder	2020	TEA	R	PLN	ADOS	E	88	S	RNN BERT	-
Learning lexical coherence representation using lstm forget gate for children with autism spectrum disorder during story-telling	2018	TEA	C	PLN	STDB	ST	67	S	SVM RNN	92,00%
Multilingual word embeddings for the assessment of narrative speech in mild cognitive impairment	2019	EA	C	PLN	Gothenburg Karolinska Dementia-Bank	CT	279	S	-	63,00%
Predicting early indicators of cognitive decline from verbal utterances	2021	EA	C	PLN	DementiaBankCT		310	S	SVM BERT Otros	81,30%
Speech pause distribution as an early marker for alzheimer's disease	2022	EA	R	PLN	Propio	E	112	S		-
Spontaneous language analysis in alzheimer's disease: evaluation of natural language processing technique for analyzing lexical performance	2021	EA	C	PLN	AD2019	CT	68	S	BERT	97,77%

Tabla 2: Resumen del estado del arte en predicción de TDAH mediante neuroimagen y de otras enfermedades mediante PLN.

Título	Año	Enfermedad	C/R	Fuente de datos	Conjunto de datos	Actividad	Tamaño	Preprocesamiento	Modelo	Precisión
Transformer-based deep neural network language models for alzheimer's disease risk assessment from targeted speech.	2021	EA	C	PLN	DementiaBankCT		269	S	BERT	88,00%
Using state of the art speaker recognition and natural language processing technologies to detect alzheimer's disease and assess its severity	2020	EA	C R	PLN	ADReSSo	E	156	S	BERT	81,00%
Using verb fluency, natural language processing, and machine learning to detect alzheimer's disease	2021	EA	C	PLN	Propio	EN	45	S	RNN	76,00%