

Análisis comparativo de algoritmos y conjuntos de datos para la detección de la enfermedad de Parkinson mediante técnicas de Deep Learning

...

Jaime Hernández Ramírez
Tecnológico Nacional México/I.T.
Orizaba
Orizaba, México
M21011175@orizaba.tecnm.mx

José Luis Sanchez Cervantes
Tecnológico Nacional México/I.T.
Orizaba
Orizaba, México
jose.sc@orizaba.tecnm.mx

Giner Alor Hernández
Tecnológico Nacional México/I.T.
Orizaba
Orizaba, México
giner.ah@orizaba.tecnm.mx

Lisbeth Rodríguez Mazahua
Tecnológico Nacional México/I.T.
Orizaba
Orizaba, México
lisbeth.rm@orizaba.tecnm.mx

Nancy Aracely Cruz Ramos
Tecnológico Nacional México/I.T.
Orizaba
Orizaba, México
dci.ncruz@ito-depi.edu.mx

Resumen — La enfermedad de Parkinson, se categoriza como una enfermedad neurodegenerativa manifestada esencialmente en adultos de edad mayor, difícilmente logra detectarse en personas menores a los 50 años. Una de sus características más graves, es no contar con una cura desarrollada y requerir tratamiento continuo mediante aplicación de medicamentos especializados. Por lo anterior, a través de diversos trabajos, investigadores han identificado patrones y atributos clave, contribuyendo en la detección y seguimiento del padecimiento. Sin embargo, se carece de herramientas tecnológicas que impulsen este proceso.

El presente artículo propone el análisis de algoritmos y datasets para la detección temprana de la enfermedad de Parkinson mediante técnicas de Deep Learning.

Los principales hallazgos permiten establecer que TensorFlow en conjunto con el modelo MobileNet V2 ofrece mejores resultados de precisión con un valor promedio de 97.62% en el proceso de clasificación de la enfermedad de Parkinson.

Palabras clave — Aprendizaje profundo, Enfermedades neurodegenerativas, Parkinson

I. INTRODUCCIÓN

Se reconoce la enfermedad de Parkinson (EP), después del Alzheimer, como el segundo padecimiento neurodegenerativo más frecuente y de constante crecimiento por adultos mayores de la población. De acuerdo con la Organización Mundial de la Salud, la EP se cataloga como una enfermedad que afecta directamente al sistema nervioso del individuo que lo padece, uno de los puntos característicos graves de esta enfermedad, es que no se cuenta actualmente con una cura y se abstiene a generar tratamientos que únicamente reduzcan el malestar presentado durante un periodo prolongado [1], [2]

Por otro lado, se cuenta con investigaciones que dedican su esfuerzo en el análisis y recolección de información clínica que brindan apoyo al personal médico y pacientes para proveer seguimiento en el tratamiento. Algunos métodos de análisis investigados para la detección de la EP, se ejecutan mediante la aplicación de ejercicios de trazabilidad, fonéticos, movimiento corporal y neurológicos, determinando así datasets con grupos de variables de interés para su análisis.

Algunas de estas variables son tiempo de ejecución, precisión y estabilidad, de acuerdo con el tipo de prueba; y aunque los resultados proveen datos exitosos, hoy en día aún se carece de herramientas tecnológicas que impulsen los procesos de recolección, análisis y clasificación de datos, dando por consiguiente el propósito general del presente análisis.

Por lo anterior, el presente artículo propone un análisis de algoritmos y datasets para la detección temprana de la EP mediante técnicas de Deep Learning. El objetivo de este análisis es recolectar y analizar los comportamientos presentados por este padecimiento a través de modelos de análisis de datos.

Este trabajo está estructurado de la siguiente manera, en la sección I se aborda la problemática y propuesta general de solución, en la sección II se aborda un análisis y comparación de investigaciones relacionadas. En la sección III se describen de manera general los conjuntos de datos de entrenamiento para incorporación en un modelo de análisis. En la sección IV se presentan los algoritmos que proveen mayor afinidad con el caso de estudio analizado, seguido. En la sección V se muestra el proceso y resultados obtenidos en la aplicación de las fuentes de conocimiento presentadas. Por último, se presentan las conclusiones.

II. ESTADO DEL ARTE

A continuación, se analizan los trabajos de investigación más relacionados con algoritmos y datasets para la identificación de la EP.

En [3] se exploró el uso de un dispositivo de pluma y tableta para estudiar diferencias en el movimiento de la mano y coordinación durante la ejecución de trazos horizontales. Los experimentos mostraron un valor promedio de precisión del 91% mediante el clasificador Naive Bayes, exponiendo como ventaja principal su versatilidad de integración con nuevas métricas de estudio frente a otros clasificadores. Por su parte, en [4] se propuso el desarrollo de cuatro modelos diferentes mediante un sistema de aprendizaje en cascada, consiguiendo de los datos obtenidos una mejora de rendimiento frente a sistemas de clasificación similares en la

detección de la EP, representando el clasificador Adaboost con una ventaja de mejora de 3.3% a través del sistema de aprendizaje propuesto. En [5] se propuso un método para la detección de episodios de congelación de la marcha en pacientes con EP, para ello, se consideró la medición de inercia con datos espectrales de monitoreo, demostrando la reproducción de los estudios, destacando el clasificador Support Vector Machine (SVM) con valor de 83% de media geométrica, sin embargo, se obtuvo una mejora de 7% adicional, incorporando el modelo 1D-ConvNet. Mientras tanto, en [6] se desarrolló un algoritmo de detección capaz de segmentar los grupos de participantes en función de la distribución de carga durante la marcha, con ello, se consiguió un valor de precisión global de 95% y se demostró como los algoritmos de clasificación de alto nivel permiten obtener resultados mejores y mayor accesibilidad, posicionando la propuesta con fines de evaluación para programas de rehabilitación. Así mismo, en [7] se implementó el aprendizaje automático para el cálculo de parámetros y determinación de clasificadores a través de sensores Leap Motion, para ello, se incorporaron operaciones de máximos y mínimos a través de cuatro clasificadores de alto nivel, donde destacó el algoritmo Random Forest (RF) con promedio global de precisión en 90.6%.

En [8], se propuso un enfoque de aprendizaje para la detección de temblores en estado de reposo, los resultados compararon el rendimiento obtenido mediante clasificadores de datos en conjunto con el enfoque propuesto, situando así, los algoritmos XGB y RF en primera instancia con métricas de precisión superiores al 70% frente a otros algoritmos. Mientras tanto, en [9] se evaluó información clínica mediante métricas estandarizadas, durante el estudio, se segmentó correctamente los grupos de participantes con base en el clasificador RF, donde se obtuvo valores promedio de sensibilidad con 96.2% y 96.9% de especificidad. Desde otro punto de vista, en [10] se utilizó un dispositivo portátil que recopiló estadísticas de supervisión motriz mediante escalas estandarizadas, la información se analizó y procesó mediante tres algoritmos, con ello, se obtuvieron mejores resultados de precisión en la implementación del algoritmo RF con un promedio de 95%. No obstante, en [11] mediante el diagnóstico del lenguaje natural para la EP, se realizó un estudio comparativo entre algoritmos de clasificación, implementando técnicas de reducción de dimensionalidad, con ello, se obtuvo que los clasificadores KNN y RF, destacaron presentando un mayor rendimiento. En un enfoque distinto, M. Wodzinski [12] presentó un modelo de detección de la EP, analizando el espectrograma vocal mediante una arquitectura ResNet, los resultados demostraron estadísticas comparables con algoritmos de alto nivel propuestos en la literatura y se obtuvo un valor promedio de precisión de 91% frente a algoritmos como RF y SVM.

En [13] se propuso un diagnóstico de rendimiento entre algoritmos de clasificación de la EP, los resultados demostraron que se logró mejorar el desempeño de los algoritmos a través de un proceso de selección de características, así mismo, SVM destacó al presentar una mejora de 79.98% a 93.84% como media de precisión. Por otro lado, en [14] se realizó el estudio de la EP basado en síntomas producidos por la pérdida de neuronas dopaminérgicas, el proceso, fue dirigido por un algoritmo clasificador de muestras, donde los resultados demostraron un mejor rendimiento de precisión con el algoritmo SVM frente a demás opciones. Haq et al., [15] enfatizaron el diagnóstico

de la comparación de algoritmos entre Machine Learning y Deep Learning mediante el análisis no invasivo de la EP, los resultados lograron identificar que las metodologías de Deep Learning mostraron mejores ventajas de rendimiento durante las pruebas, considerando Deep Neural Network y SVM como principales algoritmos de clasificación. Pahuja y Nagabhushan [16], basados en técnicas de aprendizaje automático para el análisis de la EP, proporcionaron una comparación de tecnologías para la detección de casos patológicos, como hallazgo, se obtuvo que, los algoritmos Levenberg-Marquardt y SVM presentaron rendimientos superiores al 95% de precisión durante las pruebas de detección de la EP. En [17], se propuso la generación de un modelo híbrido para la detección de la EP, su finalidad fue incorporar un sistema de capacidad incremental a través de la iteración del entrenamiento con el uso simultáneo de múltiples algoritmos, durante la investigación, se reconoció las ventajas que tiene cada uno, así como también las métricas de rendimiento individual y global, representando una solución eficaz en el área.

III. PRINCIPALES DATASETS PARA LA DETECCIÓN DE LA ENFERMEDAD DE PARKINSON

En esta sección, se presenta de manera general la investigación de diversos conjuntos de datos publicados en el ámbito del estudio de la EP mediante ejercicios de trazos de mano alzada.

Durante la revisión de resultados se analizaron las investigaciones relacionadas en función de su matriz de confusión, como herramienta que provee estadísticamente el desempeño de un algoritmo de aprendizaje supervisado. Las métricas consideradas de la matriz de confusión fueron: precisión, que es el porcentaje total de elementos clasificados correctamente, sensibilidad, que representa la clasificación de elementos positivos sobre el conjunto de verdaderos positivos y especificidad, como el valor de resultados negativos del conjunto de verdaderos negativos. A continuación, se describen los datasets analizados.

Parkinson's Drawing

Conjunto de datos de imágenes compuestas de trazos en espiral y ondas, correlacionadas en el análisis compuesto de velocidad y presión de pluma de esbozo de un grupo total de veintiocho pacientes sanos y veintisiete pacientes con padecimiento de EP evaluados en la *Escala Unificada de Calificación de la EP* [18], [19].

Parkinson Disease Spiral Drawings

Base de datos de escritura manual que consiste en 62 personas con Parkinson y quince individuos sanos distribuidos en un conjunto de datos de setenta y siete instancias y siete atributos, los resultados se clasifican en dos tipos de prueba (estática y dinámica), el conjunto de datos de estudio se obtuvo mediante una tableta digital y analizado en distintas investigaciones [20].

HandPD dataset

El conjunto de datos HandPD se compone de exámenes manuscritos de dos grupos de individuos: el grupo de sanos y el grupo de pacientes, estando este último compuesto por individuos afectados por la EP. Contiene registros de noventa y dos individuos, divididos en dieciocho personas sanas y setenta y cuatro pacientes distribuidos en un conjunto de datos de 736 imágenes [21], [22].

Parkinson's Diseases Handwriting Database (PaHaW)

Consiste en la representación de un conjunto de datos con múltiples muestras de escritura a mano realizadas por treinta y siete pacientes diagnosticados con la EP y treinta y ocho individuos sanos, los cuales, fueron emparejados con base en las características de edad y sexo [23], [24].

Performance at the clock drawing test of individuals affected by Parkinson's disease and healthy subjects

Representa un estudio retrospectivo que tiene como objetivo, verificar la precisión de la prueba de dibujo del reloj, buscando segmentar entre 240 pacientes con EP idiopática y 205 sujetos sanos, de acuerdo con los resultados obtenidos y métricas analizadas por el autor [25].

Drawing lines and circles in Parkinson's Diseases: The lateralized symptoms interface with the movements of the unaffected hand

Muestra un conjunto de datos de 36 individuos afectados por la EP y asignados a dos grupos de acuerdo con la lateralización de los síntomas presentados. Simultáneamente, se presentaron 18 participantes sanos de la misma edad integrados en el estudio mediante la tarea de acoplamiento círculos-líneas [26], [27].

En la TABLA I se presenta la comparativa de los resultados de rendimiento obtenidos en la literatura por los seis principales conjuntos de datos analizados.

TABLA I COMPARATIVA DE MÉTRICAS DE LOS PRINCIPALES CONJUNTOS DE DATOS

Conjunto de datos	Atribs.	Precisión	Sensibilidad	Especificidad
Parkinson's Drawing	2	80.00%	66.67%	93.33%
Parkinson Disease Spiral Drawings	7	77.50%	80.00%	75.00%
HandPD Dataset	7	66.72%	ND	ND
PaHaW	7	97.74%	95.50%	100%
Performance at the clock drawing	4	ND	67.00%	65.00%
Drawing lines and circles	6	ND	ND	ND

*ND – No disponible

Finalmente, se identifican los datasets: Parkinson's Drawing, Parkinson Disease Spiral Drawings y HanPD Dataset; como las fuentes de datos de mayor relevancia para la identificación de la EP, considerando factores, tales como, su afinidad en el caso de estudio analizado, su amplia documentación gracias a la gran incorporación expuesta en la literatura, así como también los resultados presentados, su libre acceso y número de variables de estudio.

IV. ALGORITMOS DE INTELIGENCIA ARTIFICIAL PARA EL ANÁLISIS DE DATOS

En esta sección, se presenta la investigación de los distintos algoritmos y herramientas de Inteligencia Artificial

(IA) para el análisis de datos que presentaron mayor afinidad con las variables de estudio.

En relación con la literatura, se realizó el análisis detallado de 20 trabajos de investigación, donde los autores, proponen la incorporación de tecnologías de la información para el seguimiento prematuro de los síntomas de la EP, incorporando técnicas de estudio mediante diversas pruebas: fonéticas, neurológicas y de movimiento corporal, así como también, pruebas de escritura y trazabilidad, por mencionar algunas.

Los modelos de análisis de datos implementados, muestran gran diversidad entre los estudios, sin embargo, entre los más comunes seleccionados por los investigadores para la revisión de elementos gráficos, se encuentran: Naive Bayes, RF y SVM, teniendo como características en común, altos índices de precisión, sensibilidad y especificidad, así como también un óptimo rendimiento en la clasificación de resultados, priorizando la calidad del estudio y manteniendo como propiedad alternativa el tiempo de ejecución.

A continuación, se presenta una descripción de los métodos de análisis antes mencionados.

Random Forest

Su tendencia de adaptarse muy cercanamente a los datos de entrenamiento y su modelo de regularización dedicado para mejorar su capacidad de clasificación mediante la combinación de múltiples árboles de decisión presentan este modelo de análisis con un amplio margen de aceptabilidad en la clasificación de recursos gráficos, convirtiéndose en una opción de mayor prioridad para el caso de estudio en el análisis de la EP. RF presentó a través de la literatura, métricas de rendimiento con valores de precisión desde 63.33% hasta 90.7%, sensibilidad de 66.77% y especificidad de 93.33%, una de las variables mejor consideradas con este método, es la precisión de los trazos graficados frente al conjunto de estudio, determinado por un modelo de coordenadas bidimensional [18], [19].

Support Vector Machine

Como modelo de aprendizaje supervisado que se utiliza tanto en problemas de clasificación como de regresión, en la detección de la EP mediante instancias gráficas de estudio tiene como objetivo encontrar la función lineal con el mayor margen en la distinción de clases. En la literatura analizada demostró alcanzar valores de precisión de 77.5% y 66.72%. Del mismo modo, algunos de los atributos de mayor relación a este método de análisis de datos fueron: coordenadas de escritura, presión, ángulo, tiempo y lateralidad del trazo [20], [21].

Naive Bayes

Algoritmo clasificador probabilístico de datos en tiempo real, representa gran utilidad en el estudio de grandes dimensiones de datos gracias a su estimación de resultados de forma independiente, a través de la literatura, Naive Bayes con una media de precisión de 59.20% por si solo, no destacó como un clasificador de gran rendimiento. Sin embargo, demostró mayor versatilidad en la incorporación de técnicas y modelos adicionales que mejoran ampliamente su precisión, al igual que presentó una mejor optimización en aplicaciones con gran cantidad de variables de estudio, reduciendo así la complejidad de análisis, tales como, lateralidad, diferencia entre radios, tiempo y dirección [22] [28].

MobileNet V2

Por otro lado, *Google Research* [29] presentó un modelo de red neuronal convolucional entrenada con un conjunto de datos de más de 2 millones de parámetros y múltiples capas dedicadas en la clasificación, optimizada para mejorar el rendimiento sin perder calidad de procesamiento en dispositivos de menor rendimiento, si bien, este recurso no se implementó en la literatura durante la detección de la EP, cuenta con investigaciones de análisis y clasificación de instancias gráficas en área de deep learning con resultados de precisión superiores al 90%, por lo que representa una alternativa de gran apoyo en el análisis desarrollado para la identificación de la EP [30], [31], [32].

Seguido, como tecnologías de soporte para la incorporación de los modelos de análisis, se estudiaron bibliotecas de aprendizaje profundo, las cuales, proporcionan las funcionalidades necesarias para el procesamiento de la información. Algunas de las herramientas identificadas en la literatura son: TensorFlow, Keras, Scikit-Learn y PyTorch, descritas a continuación.

A. TensorFlow

Plataforma de código abierto que proporciona soporte en proyectos de aprendizaje automático. TensorFlow es una herramienta de gran capacidad para el caso de estudio en el análisis de la EP gracias a su variedad de herramientas y capacidad de predicción de producción a gran escala mediante procesamiento dedicado, fundamental en el estudio de datasets de gran variedad de propiedades. Así mismo, cuenta con un amplio catálogo de modelos y algoritmos, con la capacidad de acoplarse directamente en el estudio de redes convolucionales y redes neuronales recurrentes, incorporando los algoritmos en cuestión de estudio.

B. Keras

Biblioteca de código abierto eficaz en la creación de redes neuronales. Keras es una herramienta de gran potencial para el caso de estudio de la EP gracias a que ofrece una amplia compatibilidad entre plataformas y mantiene un uso óptimo de recursos. Así mismo, proporciona un amplio soporte en la creación de redes convolucionales y redes neuronales, lo que provee una amplia ventaja en la implementación de los tres algoritmos analizados previamente en la literatura. Adicionalmente, Keras mantiene soporte directo e incorporado como biblioteca de TensorFlow, impulsando su desarrollo como característica complementaria.

C. Scikit-Learn

Biblioteca de código abierto enfocada en la implementación fácil y rápida de modelos de aprendizaje automático en el lenguaje de programación Python, se presenta como herramienta de gran afinidad en el estudio de la EP, gracias a su directa y amplia integración con los modelos de clasificación de instancias, así como también su optimización de procesos de análisis que mejora la eficiencia de los algoritmos mediante la integración de código en C a través de la biblioteca Cython y soporte de variaciones en los métodos SVM.

D. PyTorch

Biblioteca de código abierto dedicada en la programación optimizada de tensores. PyTorch se destaca en el caso de estudio de análisis de la EP, gracias a su fácil implementación y relación directa con el lenguaje de programación Python, mismo, que es considerado como uno de los lenguajes mejor

considerados en el área de procesamiento de información para la IA. PyTorch integra la representación de gráficos dinámicos para la realización de cálculos, mejorando la depuración de código y puesta en marcha, óptimo en el análisis de grandes conjuntos de datos encontrados en la literatura. PyTorch se presenta como una herramienta mejorada a la plataforma de datos TensorFlow y no cuenta con integración directa en la nube, por lo que requiere de la incorporación de recursos de hardware o plataformas de terceros para su implementación.

V. CASO DE ESTUDIO

En esta sección, se presenta un caso de estudio realizado durante el desarrollo de la investigación, mostrando el proceso realizado, las pruebas contempladas y los resultados obtenidos por los algoritmos de análisis, conjuntos de datos de entrenamiento, plataformas de integración y rendimiento obtenido en el contexto mencionado.

Como punto de partida, se preparó el modelo de datos que utilizó el algoritmo como fuente de información para su entrenamiento y pruebas de ejecución, para ello, se tomó como referencia la combinación de dos de los seis conjuntos de datos analizados en la sección III ([18], [20]), el parámetro de selección fue el modelo de instancias recopiladas durante la investigación, correspondiente a un modelo de trazo en espiral para el caso de estudio, con ello, se complementó un nuevo repositorio con un total de 216 instancias generadas por 117 participantes segmentados en 2 grupos.

El modelo de equivalencia, consistió en la segmentación de 108 muestras de participantes diagnosticados con la EP y 108 muestras de participantes saludables, posteriormente, se realizó una segmentación adicional correspondiente al 80% de instancias como conjunto de entrenamiento, excluyendo el 20% como conjunto de validación, conservando un total de 172 instancias para entrenamiento.

Dicho lo anterior, el siguiente proceso se enfocó en el desarrollo de los modelos de clasificación, con el fin de preparar e implementar el algoritmo con el repositorio de datos generado, los candidatos seleccionados en cuestión para el caso de prueba, fueron referentes a los dos algoritmos de mayor relevancia a través de la literatura por factor de rendimiento, RF y SVM, adicionalmente, se decidió incorporar en la comparación, el algoritmo MobileNet V2 (MN2) debido a su gran rendimiento en pruebas de clasificación.

Como tecnologías de soporte para la implementación de los modelos, se definió la selección por factor de relevancia en la literatura, donde se expresó como mejor resultado las plataformas TensorFlow y Scikit-Learn como opciones de implementación para el caso de estudio a analizar. Los resultados obtenidos por los procesos de entrenamiento y verificación se visualizan a continuación, en la TABLA II se representan los resultados obtenidos por la plataforma Scikit-Learn, mientras que en la TABLA III se exponen los datos recuperados por la plataforma TensorFlow.

TABLA II RENDIMIENTO DE LOS MODELOS CON SCIKIT-LEARN

		Algoritmo		
		SVM	RF	MN2
Métricas	Precisión	73.33%	83.33%	96.56%
	Sensibilidad	80.00%	73.33%	95.23%
	Especificidad	66.67%	93.33%	97.97%

TABLA III RENDIMIENTO DE LOS MODELOS EN TENSORFLOW

		Algoritmo		
		SVM	RF	MN2
Métricas	Precisión	73.52%	81.86%	97.62%
	Sensibilidad	71.42%	78.20%	92.86%
	Especificidad	76.08%	86.51%	92.86%

Acorde con lo anterior, se observó sobre la comparación de rendimiento en los algoritmos seleccionados por la literatura, RF destacó notablemente en su implementación con la plataforma Scikit-Learn, donde obtuvo un valor promedio de precisión de 83.33%, en cambio, MN2, se posicionó como la solución de mayor rendimiento en su aplicación desde ambas plataformas, sin embargo, obtuvo un mejor rendimiento en su implementación dentro de la plataforma TensorFlow con un valor promedio de 97.62%.

En segunda instancia, para continuar con la implementación y prueba de los modelos de clasificación en el presente caso de estudio, se desarrolló un módulo de software, cuyo propósito, es proporcionar al usuario la funcionalidad de clasificar muestras de trazos digitales obtenidos por los pacientes mediante una plataforma de análisis, debido a que la variación de rendimiento con el algoritmo MN2 es mínima entre ambas plataformas, para el caso práctico, se decidió implementar la plataforma Scikit-Learn con el algoritmo RF, mientras que, por otro lado, la plataforma TensorFlow se estudió con el algoritmo MN2.

Como instancias de estudio, se obtuvieron adicionalmente muestras de diez personas de prueba, cinco de ellas diagnosticadas médicamente con la EP y cinco más, diagnosticadas como personas sanas, solicitando que realizaran una serie de seis ejercicios de trazabilidad, dibujando espirales guiados por una plantilla. En la FIGURA 1 se muestran algunos de los resultados obtenidos por las personas con padecimiento confirmado, mientras que en la FIGURA 2, se presentan ejemplos de aquellos modelos obtenidos por personas sanas.

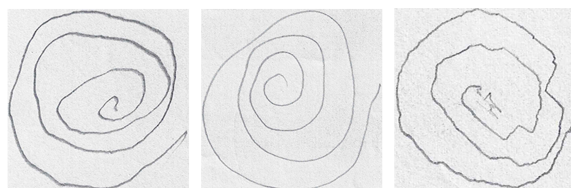


FIGURA 1 MUESTRAS PARA ANÁLISIS POSITIVAS

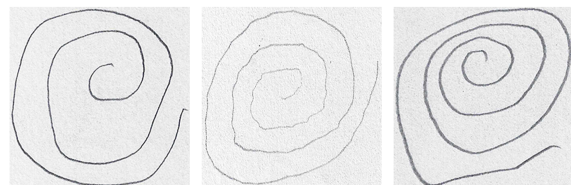


FIGURA 2 MUESTRAS PARA ANÁLISIS NEGATIVAS

Posteriormente, para el caso práctico de integración, el módulo desarrollado se compone de dos interfaces gráficas principales, descritas a continuación:

Con respecto al análisis, el usuario cuenta con una interfaz (FIGURA 3) donde realiza la carga de una muestra de trazabilidad gráfica obtenida por el paciente, así mismo, permite la selección de la biblioteca de análisis de preferencia, dato que permite comparar los resultados obtenidos entre las distintas plataformas propuestas, una vez adjunta la muestra y seleccionada la biblioteca de estudio, los datos se envían para su análisis.

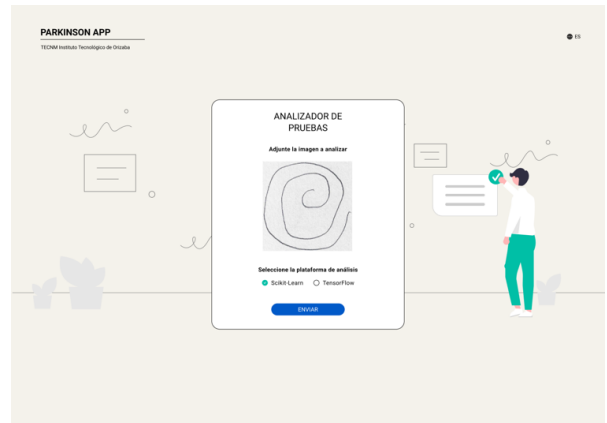


FIGURA 3 INTERFAZ DE CARGA DE PRUEBAS DE EP

Finalizado el proceso de análisis, el módulo muestra en pantalla los resultados obtenidos, permitiendo al usuario visualizar cada una de las métricas obtenidas, así como también, la previsualización del estudio enviado, el resultado global que indica la posibilidad del padecimiento y la plataforma de análisis seleccionada, en la FIGURA 4, se muestra un ejemplo detallado.

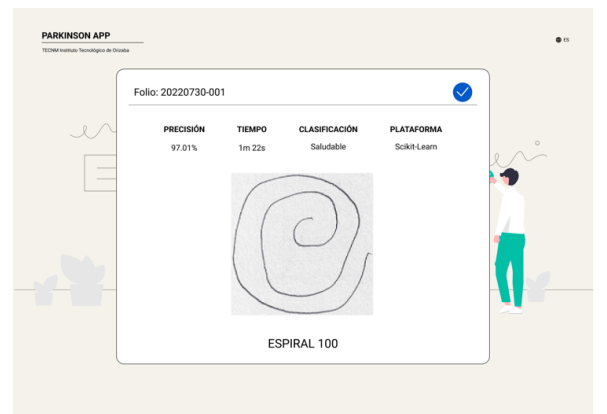


FIGURA 4 INTERFAZ DE RESULTADOS DEL ANÁLISIS DE PRUEBAS DE LA EP

La primera prueba consistió en adjuntar de forma aleatoria cada una de las muestras obtenidas de los pacientes, seleccionando como plataforma de análisis Scikit-Learn, los resultados obtenidos demostraron clasificar correctamente el 91.66% de los casos presentados en un tiempo de ejecución de 82 segundos.

Para la segunda prueba realizada, se seleccionó el uso de la plataforma TensorFlow, para este caso, los resultados

demonstraron que el clasificador logró segmentar en su totalidad el conjunto de datos enviado de manera correcta, sin embargo, el tiempo de ejecución presentado fue mayor, con un total de 155 segundos.

En la TABLA IV, se muestran a manera de resumen los resultados obtenidos durante el caso práctico presentado.

TABLA IV RESULTADOS DE ANÁLISIS DE LAS PRUEBAS DURANTE EL CASO PRÁCTICO

Plataforma	Algoritmo	Aciertos	Tiempo
Scikit-Learn	RF	91.66%	82 seg
TensorFlow	MN2	100.00%	155 seg

Para comprender mejor el análisis del rendimiento obtenido en cada plataforma de estudio, en la TABLA V, se muestra el resultado de rendimiento de los algoritmos, incorporando la métrica de tiempo sobre el número de instancias utilizadas para el caso práctico.

TABLA V RENDIMIENTO DE LOS MODELOS DE CLASIFICACIÓN

		Plataforma	
		Scikit-Learn (RF)	TensorFlow (MN2)
Métricas	Precisión	83.33%	97.62%
	Sensibilidad	73.33%	92.86%
	Especificidad	93.33%	92.86%
	Tiempo (60 instancias)	82 seg	155 seg

Para simplificar la comparativa realizada, en la FIGURA 5, se expone un gráfico comparativo de los promedios obtenidos de cada una de las bibliotecas de soporte bajo las cuales se realizaron los análisis de prueba.

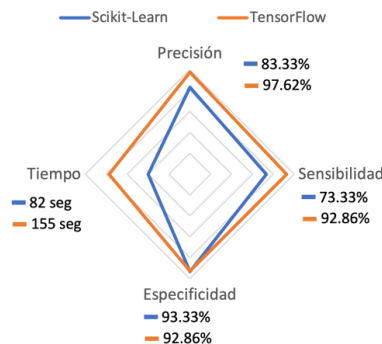


FIGURA 5 COMPARACIÓN DE RESULTADOS ENTRE BIBLIOTECAS SCIKIT-LEARN Y TENSORFLOW

CONCLUSIONES

El constante incremento de pacientes con EP es un tema muy alarmante, ya que al no contar con herramientas que apoyen a la detección temprana, dificulta el tratamiento oportuno de los pacientes, es por ello que, en la actualidad, se apuesta en la inversión de tiempo y esfuerzo en el desarrollo e implementación de nuevas herramientas que permitan de

manera oportuna, detectar a personas propensas a padecer esta enfermedad.

En el análisis presentado a través de la investigación, se observa la existencia de diversos trabajos e investigaciones realizadas que apuestan en la implementación de técnicas de aprendizaje automático y aprendizaje profundo, con el fin de apoyar al personal médico en la detección temprana de la enfermedad mediante el análisis de estudios clínicos y comportamiento de los pacientes.

Así mismo, tras haber realizado un análisis minucioso de los distintos datasets propuestos, al igual que los diversos métodos de análisis de datos mayormente utilizados en la literatura y haberle puesto a prueba mediante un caso de uso específico, se demostró que los mejores resultados para la clasificación de instancias de estudios de la EP, se obtuvieron con la biblioteca TensorFlow. TensorFlow demostró contar con mejores métricas de estudio frente a demás bibliotecas y métodos analizados, obteniendo así un valor promedio de precisión de 97.62% en un tiempo máximo de 155 segundos durante las pruebas y 100 ms de ejecución individual. Con base en lo anterior, los resultados obtenidos son lo suficientemente acertados para implementar estas técnicas como método auxiliar en la detección y seguimiento en el tratamiento de la EP.

Como trabajo a futuro, se expandirá la capacidad del modelo presentado con mayor rendimiento para realizar estudios de detección de la EP a través de nuevos sistemas de información, tales como análisis de datos del lenguaje natural y supervisión de la actividad motriz. Por otro lado, el modelo presentado, mantendrá el acceso como una plataforma enfocada a la identificación, diagnóstico, monitorización y tratamiento de enfermedades cronicodegenerativas y neurodegenerativas.

Considerando la problemática presentada y con la finalidad de ayudar en la detección de este padecimiento, se promueve la integración de tecnologías de la información para la detección temprana de la EP mediante técnicas de Deep Learning, donde se presente al paciente una herramienta para la generación de trazos y con ayuda de conjuntos de datos de estudios previamente efectuados, se analicen los resultados, notificando así, posibles casos de padecer la EP.

AGRADECIMIENTOS

Este trabajo fue patrocinado por el Consejo Nacional de Ciencia y Tecnología (CONACYT), el Tecnológico Nacional de México (TecNM) y la Secretaría de Educación Pública (SEP).

REFERENCIAS

- [1] "Enfermedad de Parkinson," *Organización Mundial de la Salud*, Jun. 13, 2022. <https://www.who.int/es/news-room/fact-sheets/detail/parkinson-disease> (accessed Jul. 21, 2022).
- [2] "El Parkinson, la segunda enfermedad neurodegenerativa más común," *Fundación Cien*, Apr. 16, 2018. <https://www.fundacioncien.es/noticias/el-parkinson-la-segunda-enfermedad-neurodegenerativa-mas-comun> (accessed Jul. 21, 2022).
- [3] C. Kotsavasiloglou, N. Kostikis, D. Hristu-Varsakelis, and M. Arnaoutoglou, "Machine learning-based classification of simple drawing movements in Parkinson's disease," *Biomed Signal Process Control*, vol. 31, pp. 174–180, Jan. 2017, doi: 10.1016/j.bspc.2016.08.003.

- [4] L. Ali, C. Zhu, N. A. Golilarz, A. Javeed, M. Zhou, and Y. Liu, "Reliable Parkinson's Disease Detection by Analyzing Handwritten Drawings: Construction of an Unbiased Cascaded Learning System Based on Feature Selection and Adaptive Boosting Model," *IEEE Access*, vol. 7, pp. 116480–116489, 2019, doi: 10.1109/ACCESS.2019.2932037.
- [5] J. Camps *et al.*, "Deep learning for freezing of gait detection in Parkinson's disease patients in their homes using a waist-worn inertial measurement unit," *Knowl Based Syst*, vol. 139, pp. 119–131, 2018, doi: 10.1016/j.knsys.2017.10.017.
- [6] R. Alkhatib, M. O. Dlab, C. Corbier, and M. el Badaoui, "Machine Learning Algorithm for Gait Analysis and Classification on Early Detection of Parkinson," *IEEE Sens Lett*, vol. 4, no. 6, Jun. 2020, doi: 10.1109/LESENS.2020.2994938.
- [7] A. Moshkova, A. Samorodov, N. Voinova, E. Ivanova, and E. Fedotova, "Parkinson's Disease Detection by Using Machine Learning Algorithms and Hand Movement Signal from LeapMotion Sensor," 2020.
- [8] Lin Yao, Peter Brown, and Mahsa Shoaran, *Resting Tremor Detection in Parkinson's Disease with Machine Learning and Kalman Filtering*. 2018.
- [9] S. Arora *et al.*, "Detecting and monitoring the symptoms of Parkinson's disease using smartphones: A pilot study," *Parkinsonism Relat Disord*, vol. 21, no. 6, pp. 650–653, Jun. 2015, doi: 10.1016/j.parkreldis.2015.02.026.
- [10] E. Rovini, C. Maremmanni, A. Moschetti, D. Esposito, and F. Cavallo, "Comparative Motor Pre-clinical Assessment in Parkinson's Disease Using Supervised Machine Learning Approaches," *Ann Biomed Eng*, vol. 46, no. 12, pp. 2057–2068, Dec. 2018, doi: 10.1007/s10439-018-2104-9.
- [11] A. Anand, A. Haque, J. Sahaya, R. Alex, and N. Venkatesan, "Evaluation of Machine learning and Deep learning algorithms combined with dimensionality reduction techniques for classification of Parkinson's Disease."
- [12] Marek Wodzinski, Andrzej Skalski, Daria Hemmerling, Juan Rafael Orozco-Arroyave, and Elmar N'oth, *Deep Learning Approach to Parkinson's Disease Detection Using Voice Recordings and Convolutional Neural Network Dedicated to Image Classification*. 2019.
- [13] Z. Karapinar Senturk, "Early diagnosis of Parkinson's disease using machine learning algorithms," *Med Hypotheses*, vol. 138, May 2020, doi: 10.1016/j.mehy.2020.109603.
- [14] R. Prashanth, S. Dutta Roy, P. K. Mandal, and S. Ghosh, "High-Accuracy Detection of Early Parkinson's Disease through Multimodal Features and Machine Learning," *Int J Med Inform*, vol. 90, pp. 13–21, Jun. 2016, doi: 10.1016/j.ijmedinf.2016.03.001.
- [15] A. U. Haq *et al.*, "Comparative Analysis of the Classification Performance of Machine Learning Classifiers and Deep Neural Network Classifier for Prediction of Parkinson Disease," 2018.
- [16] G. Pahuja and T. N. Nagabhushan, "A Comparative Study of Existing Machine Learning Approaches for Parkinson's Disease Detection," *IETE Journal of Research*, vol. 67, no. 1. Taylor and Francis Ltd., pp. 4–14, 2021. doi: 10.1080/03772063.2018.1531730.
- [17] M. Nilashi, O. Ibrahim, H. Ahmadi, L. Shahmoradi, and M. Farahmand, "A hybrid intelligent system for the prediction of Parkinson's Disease progression using machine learning techniques," *Biocybern Biomed Eng*, vol. 38, no. 1, pp. 1–15, 2018, doi: 10.1016/j.bbe.2017.09.002.
- [18] Scot K, "Parkinson's Drawings," 2019. <https://www.kaggle.com/datasets/kmader/parkinsons-drawings> (accessed Jun. 23, 2022).
- [19] P. Zham, D. K. Kumar, P. Dabnichki, S. Poosapadi Arjunan, and S. Raghav, "Distinguishing Different Stages of Parkinson's Disease Using Composite Index of Speed and Pen-Pressure of Sketching a Spiral," *Front Neurol*, vol. 8, Sep. 2017, doi: 10.3389/fneur.2017.00435.
- [20] Team AI, "Parkinson Disease Spiral Drawings," 2017. <https://www.kaggle.com/datasets/team-ai/parkinson-disease-spiral-drawings> (accessed Jun. 23, 2022).
- [21] Pereira C, "HandPD dataset," 2020. <https://www.kaggle.com/datasets/claytonteybauru/spiral-handpd> (accessed Jun. 23, 2022).
- [22] C. R. Pereira *et al.*, "A new computer vision-based approach to aid the diagnosis of Parkinson's disease," *Comput Methods Programs Biomed*, vol. 136, pp. 79–88, Nov. 2016, doi: 10.1016/j.cmpb.2016.08.005.
- [23] BDALab, "Parkinson's Disease Handwriting Database (PaHaW)," 2020. <https://bdalab.utko.fekt.vut.cz> (accessed Jun. 23, 2022).
- [24] J. Mucha *et al.*, "Identification and Monitoring of Parkinson's Disease Dysgraphia Based on Fractional-Order Derivatives of Online Handwriting," *Applied Sciences*, vol. 8, no. 12, p. 2566, Dec. 2018, doi: 10.3390/app8122566.
- [25] F. Scarpina, C. Paschino, L. Priano, and A. Mauro, "Performance at the clock drawing test of individuals affected by Parkinson's disease and healthy subjects: a retrospective study," *Neurological Sciences*, vol. 41, no. 4, pp. 843–849, Apr. 2020, doi: 10.1007/s10072-019-04167-w.
- [26] F. Scarpina *et al.*, "Drawing lines and circles in Parkinson's Disease: The lateralized symptoms interfere with the movements of the unaffected hand," *Neuropsychologia*, vol. 151, p. 107718, Jan. 2021, doi: 10.1016/j.neuropsychologia.2020.107718.
- [27] Scarpina F, "Drawing lines and circles in Parkinson's Disease: The lateralized symptoms interfere with the movements of the unaffected hand," Jan. 2021. <https://zenodo.org/record/4429668#.YrWESS-xBQI> (accessed Jun. 23, 2022).
- [28] J. M. Hausdorff, A. Lertratanakul, M. E. Cudkovicz, A. L. Peterson, D. Kaliton, and A. L. Goldberger, "Dynamic markers of altered gait rhythm in amyotrophic lateral sclerosis," *J Appl Physiol*, vol. 88, no. 6, pp. 2045–2053, Jun. 2000, doi: 10.1152/jappl.2000.88.6.2045.
- [29] Sandler Mark, "MobileNetV2: The Next Generation of On-Device Computer Vision Networks," *Google AI Blog*, Apr. 2018.
- [30] K.-W. Lee, H.-J. Lee, H. Hu, and H.-J. Kim, "Analysis of facial ultrasonography images based on deep learning," *Sci Rep*, vol. 12, no. 1, p. 16480, Oct. 2022, doi: 10.1038/s41598-022-20969-z.
- [31] S. A. Sanjaya and S. Adi Rakhmawan, "Face Mask Detection Using MobileNetV2 in The Era of COVID-19 Pandemic," in *2020 International Conference on Data Analytics for Business and Industry: Way Towards a Sustainable Economy (ICDABI)*, Oct. 2020, pp. 1–5. doi: 10.1109/ICDABI51230.2020.9325631.
- [32] C. H. Karadal, M. C. Kaya, T. Tuncer, S. Dogan, and U. R. Acharya, "Automated classification of remote sensing images using multileveled MobileNetV2 and DWT techniques," *Expert Syst Appl*, vol. 185, p. 115659, Dec. 2021, doi: 10.1016/j.eswa.2021.115659.