

Reconocimiento Continuo de Actividades de la Vida Diaria en Pacientes con Temblor Patológico

A. Ruiz-Vitte¹, A. Comesaña², B. Larraga¹, E. Rocon², A. Gutiérrez¹

¹ ETSI Telecomunicación, Universidad Politécnica de Madrid, Madrid, España.

² Centro de Automática y Robótica, CSIC-UPM, Madrid, España

Resumen

El temblor patológico, que afecta la capacidad de realizar actividades diarias, presenta desafíos para las estrategias terapéuticas actuales basadas en la evaluación clínica subjetiva y el manejo sintomático. Este artículo introduce un sistema innovador para la clasificación de actividades en pacientes con temblor, utilizando modelos entrenados con datos de series temporales segmentadas en diferentes tamaños de ventana. El sistema logró una precisión global del 73%, aunque el rendimiento varió entre las tareas. Esto supone un avance en el estado actual de los sistemas de clasificación de actividades en tiempo real, mostrando el potencial para mejorar la monitorización del temblor y optimizar la dosis de medicación.

1. Introducción

El temblor patológico, caracterizado por oscilaciones involuntarias y rítmicas que afectan principalmente a las manos, es el trastorno del movimiento más frecuente en la práctica médica (1). El temblor esencial (TE) afecta a alrededor del 4% de la población mayor de 65 años, mientras que la enfermedad de Parkinson (EP) es el segundo trastorno neurodegenerativo más frecuente después de la enfermedad de Alzheimer (2,3). Estos síntomas afectan profundamente a la calidad de vida de los pacientes, mermando su capacidad para realizar actividades de la vida diaria (*Activities Of Daily Living* en inglés, ADLs) desde las primeras fases de la enfermedad y a lo largo de su progresión (4).

Aunque la medicación es el tratamiento más eficaz para el temblor, a menudo se prescribe mediante un proceso de ensayo y error, lo que puede llevar a una amplia gama de efectos secundarios, representando un importante desafío en las estrategias terapéuticas actuales (5). Los síntomas motores derivados del temblor patológico suelen evaluarse mediante demostraciones mecánicas del temblor y cuantificarse utilizando escalas clínicas. Sin embargo, estos métodos están limitados por el sesgo de rendimiento del paciente, influido por efectos como el placebo o el 'síndrome de la bata blanca' (4,6), en el que los pacientes realizan un esfuerzo adicional en presencia del clínico, lo que conduce a una representación inexacta de sus capacidades motoras. Este síndrome también dificulta la evaluación de la eficacia de la medicación, lo cual es esencial para estudiar la evolución de los síntomas con los ajustes de dosis (7).

En un intento por minimizar la subjetividad que conlleva la evaluación de los trastornos del movimiento, se hace uso de unidades de medición inercial (*Inertial Measurement Units* en inglés, IMUs). Proporcionar información sobre las tareas

realizadas durante los episodios de temblor podría mejorar la comprensión del estado clínico del paciente. Sin embargo, la aplicación de estos métodos de identificación de tareas a pacientes que presentan temblor no ha arrojado resultados concluyentes (7–9). A menudo, los estudios en los que participan pacientes no suelen abarcar la identificación de tareas cotidianas complejas, centrándose en cambio en actividades posturales y de transición, y los que sí abordan la identificación de tareas complejas utilizan múltiples IMUs, lo que puede resultar incómodo para los pacientes.

En cuanto a la identificación de tareas, la selección precisa del tamaño de la ventana es crucial para una clasificación eficaz de las actividades. Sin embargo, este aspecto sigue siendo ambiguo e indefinido (10–12), y la mayoría de los diseños se basan en estimaciones de estudios anteriores sin un respaldo empírico significativo.

La principal contribución de este trabajo es la clasificación de actividades complejas a partir de secuencias continuas de ADLs, introduciendo una segmentación variable para la extracción de características y abordando la falta de consenso sobre el tamaño óptimo de ventana. Para la recolección de las señales, se utiliza una pulsera smartwatch que registra los datos de movimiento de la muñeca en tiempo real. Este enfoque permitirá a los neurólogos mejorar la monitorización del temblor y evaluar el impacto de la medicación en la calidad de vida de los pacientes durante las ADLs, con el objetivo de estudiar la evolución de la enfermedad mediante una monitorización continua.

2. Metodología

2.1 Recopilación de Datos

Basándose en estudios anteriores (13) en los que se introduce brevemente el enfoque presentado en este trabajo, se segmenta una población de 70 sujetos: 63 de ellos (5 TE, 5 EP, 37 participantes de control) y un conjunto adicional de datos de 16 pacientes de temblor de tipo indeterminado (4,14) se utilizaron para entrenar y testear el clasificador de ADLs, mientras que de los restantes (7 participantes de control) se extrajeron señales continuas para validar la aplicación del clasificador en un entorno doméstico diario, donde llevaban la pulsera durante períodos prolongados de tiempo.

En un primer paso, se utilizó una pulsera smartwatch (Fitbit Sense, (15)) que se colocó en la muñeca dominante para recoger los datos, registrando la aceleración lineal y la velocidad angular en tres ejes (x, y, z). Los datos de

movimiento se muestrearon a 30 Hz. A continuación, todos los datos se enviaron a un dispositivo Android que los transfirió a una base de datos para su almacenamiento. Durante las pruebas para el entrenamiento del modelo, cada participante realizó nueve tareas de la vida diaria a una velocidad elegida por él mismo, de la misma forma en que las realizaría normalmente:

- Peinarse (CH)
- Abrocharse los botones (BB)
- Cortar con cuchillo y comer con tenedor los trozos previamente cortados (EF)
- Simular beber (SD)
- Abrir y cerrar una caja (OB)
- Pasar páginas (TP)
- Escribir (WW)
- Cepillarse los dientes (BT)
- Abrir y/o cerrar una puerta (TD)

Cada participante del grupo de entrenamiento realizó estas tareas, repitiendo cada una de ellas entre tres y seis veces. El movimiento entre repeticiones y la realización de tareas también se registró y etiquetó como 'no tarea' (NT).

En una segunda fase, para obtener señales continuas del grupo de validación en el entorno doméstico, los sujetos llevaron el smartwatch durante secuencias de tiempo prolongado, marcando en el dispositivo Android cuándo empezaban o terminaban cada tarea, sin ningún orden predeterminado.

El estudio siguió la Declaración de Helsinki y contó con el consentimiento informado de los participantes para los ensayos de usabilidad del sistema de monitorización del movimiento del proyecto NETremor (TED2021-130174B-C31).

2.1 Preprocesamiento y Filtrado

En primer lugar, se realizó un cribado y limpieza de datos para descartar señales fuera del centro de distribución y evitar grabaciones defectuosas. Siguiendo el procesamiento del conjunto adicional de datos, se aplicó un filtro de Butterworth con una frecuencia de corte de 4 Hz y orden 6 para separar los movimientos voluntarios del temblor, garantizando que el temblor se filtrara sin afectar al movimiento voluntario (14).

2.2 Extracción de Características

Basándose en las conclusiones derivadas de estudios anteriores (13), las señales se dividieron en ventanas superpuestas de diferentes tamaños. Las ventanas seleccionadas fueron de 5, 10, 15 y 20 segundos con 2, 4, 6 y 8 segundos de solapamiento respectivamente.

Se extrajeron 15 características temporales y espectrales por cada ventana, que fueron seleccionadas en base a estudios previos (4,13,16). La extracción de características de cada tipo de ventana dio lugar a cuatro colecciones separadas de 90 características, correspondientes a los cuatro procesos de segmentación realizados. Para hacer frente a la complejidad y el tamaño del conjunto de datos, se utilizó *SelectKBest* (17) para seleccionar las características con las puntuaciones más altas e incluirlas en el subconjunto final

de características. Estas fueron por una parte características básicas de dispersión y variabilidad (promedio, desviación estándar, máximo, mínimo y varianza) en ambos sensores, y por otra, características de forma y complejidad (asimetría, curtosis, entropía y centroide) de forma más limitada, especialmente en el giroscopio, donde el eje x presenta menos características de complejidad, y con mayor presencia en las medidas del acelerómetro, especialmente en los ejes x y z.

2.3 Estrategia de clasificación

Aplicamos diferentes tamaños de ventana para clasificar señales continuas y de mayor duración de sujetos sin temblor (13). Este enfoque se apoya en la idea de variar el tamaño de las ventanas puede optimizar la clasificación de tareas y proporcionar una visión holística para el análisis de señales. Se entrenó un conjunto de modelos basados en máquinas de vector de soporte (*Support Vector Machine* en inglés, SVM) utilizando los datos del grupo de los 63 sujetos. El modelo empleó un enfoque algorítmico en dos fases (véase Figura 1):

En la primera fase, los datos se segmentaron en conjuntos de entrenamiento y de prueba utilizando cuatro tamaños de ventana diferentes. Con estos segmentos de ventana se entrenaron cuatro clasificadores distintos. Para cada secuencia de tareas, se asignó una etiqueta basada en la moda de etiquetas predichas o, en casos de ambigüedad, la etiqueta con la probabilidad media más alta. Este proceso garantizaba que cada secuencia se etiquetara con una probabilidad de predicción asociada.

En la segunda fase, se agregaron las predicciones de todos los clasificadores. La etiqueta final de la tarea y la probabilidad de predicción para cada secuencia se determinaron repitiendo el mismo procedimiento que en la fase anterior, seleccionando la moda o la etiqueta de máxima probabilidad en caso de ambigüedad.

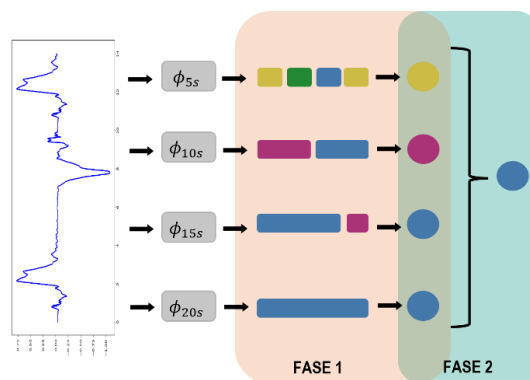


Figura 1: Esquema del algoritmo utilizado para la clasificación.

Se utilizó el clasificador SVM (18) por su rendimiento superior en pruebas anteriores (13). Se optimizó mediante ajuste de hiperparámetros con *GridSearch* (17) y se validó con validación cruzada de 7 particiones, manteniendo la proporcionalidad de las clases. Los datos se dividieron en un 85% para entrenamiento y 15% para prueba. Una vez entrenado, se usaron las señales preprocesadas de larga duración del grupo de validación para identificar y clasificar tareas en secuencias continuas.

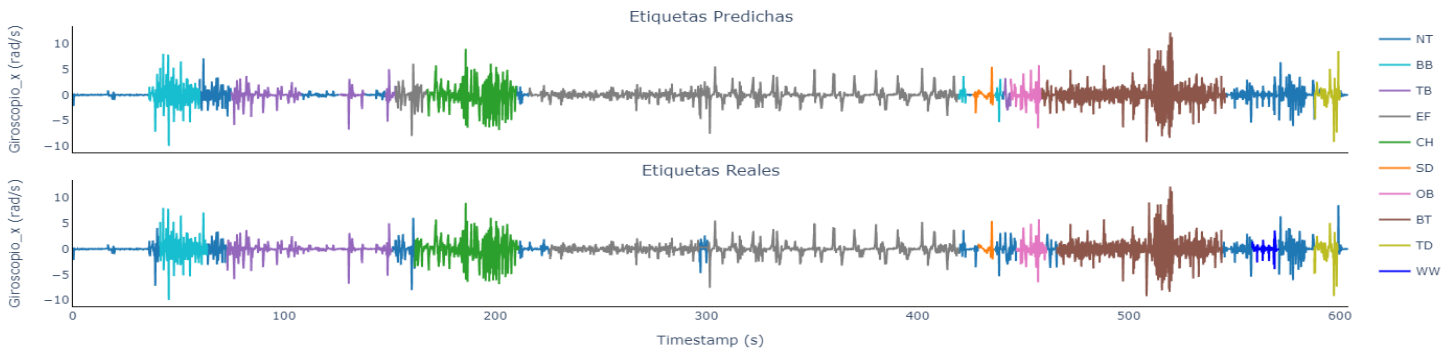


Figura 2: Gráficos que comparan las etiquetas predichas con las verdaderas en una secuencia continua de ADLs.

3. Resultados

Para mostrar el rendimiento del modelo, se construyen gráficos como el de la Figura 1 para cada paciente, en el que diferentes colores representan distintas tareas a lo largo del tiempo, incluida la ‘no tarea’, que denota la ausencia de cualquier tarea específica. En la imagen superior se muestran las tareas realizadas, mientras que en la inferior las clasificadas con el modelo.

La Tabla 1 presenta las métricas que evalúan el rendimiento del modelo. La precisión global es del 72,77%. Sin embargo, existe una notable variabilidad entre ellas

Tareas	Precisión	Recall	F1-score
BB	26,64%	79,74%	39,93%
BT	66,32%	88,31%	75,75%
CH	50,34%	85,68%	63,42%
EF	72,75%	58,84%	65,06%
NT	85,59%	77,65%	81,43%
OB	29,63%	54,36%	38,35%
SD	57,40%	48,57%	52,62%
WW	59,77%	58,05%	58,90%
TB	56,13%	76,23%	64,66%
TD	62,42%	21,06%	31,50%
Accuracy	-	-	72,77%

Tabla 1: Resultados obtenidos por cada tarea.

La matriz de confusión de la Figura 3 ilustra el rendimiento del modelo en las distintas clases. Aunque la matriz muestra una diagonal fuerte, lo que indica predicciones precisas para muchas tareas, también revela errores de clasificación significativos, especialmente falsos positivos con la tarea NT. Esto sugiere que, aunque el modelo es generalmente eficaz, existen áreas específicas donde su capacidad para distinguir entre tareas podría mejorarse.

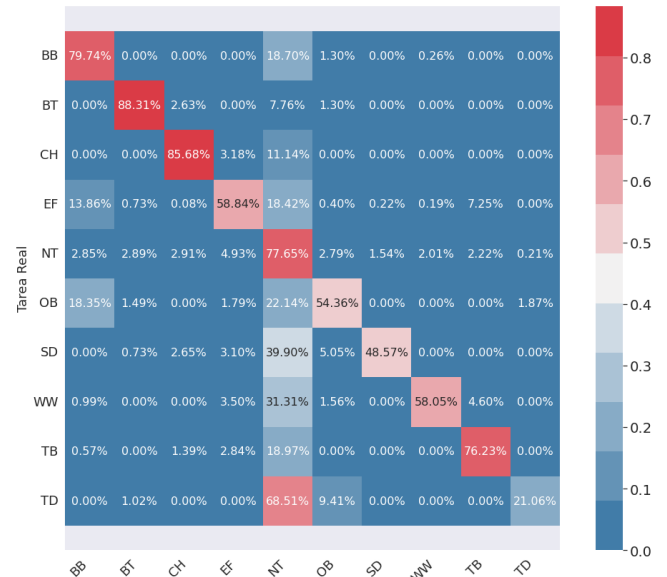


Figura 3: Matriz de confusión de la clasificación de ADLs.

4. Discusión

Observando la Figura 2, se observa que el sistema es capaz de clasificar correctamente las tareas dentro de un margen de error. Se presenta, por lo tanto, un enfoque innovador que aborda la discrepancia en el tamaño de ventana, permitiendo la clasificación de tareas complejas mediante el uso de un único sensor, representando una contribución significativa al estado del arte. La evaluación del modelo revela que, a pesar de una precisión global del 73%, persisten desafíos significativos en la clasificación de ADLs en señales prolongadas.

Al analizar la Figura 1, se observa que, aunque el sistema identifica correctamente la mayoría de las tareas, persisten algunos errores de clasificación. Por ejemplo, se detectan casos tempranos de OB y señales muy cortas de BB, que no son coherentes con el comportamiento esperado. Esta incoherencia sugiere que se podrían implementar mecanismos de postprocesamiento para corregir clasificaciones potencialmente erróneas basadas en información contextual.

Esta falta de precisión se refleja en las métricas presentadas en la Tabla 1. Para tareas como BB y OB, potencialmente más complejas debido a su carácter bimanual, la precisión

es baja (< 30%), a pesar de una alta exhaustividad (*recall* en inglés), lo que indica la presencia de falsos positivos, como se observa en la Figura 3. En contraste, tareas como BT y CH muestran métricas de rendimiento equilibradas, lo que sugiere una clasificación más confiable. La tarea con las métricas más altas es NT, probablemente debido a su mayor representación en el conjunto de datos. La matriz de confusión respalda esta observación, revelando que muchas tareas se clasifican erróneamente como NT, probablemente debido al desequilibrio de clases.

Estos resultados sugieren que, aunque el sistema es prometedor, especialmente en ciertas tareas, es necesario refinarlo para abordar los problemas relacionados con el desequilibrio de clases y mejorar la capacidad del modelo para distinguir entre tareas estrechamente relacionadas o bimanuales. La implementación de técnicas de posprocesamiento y estrategias para mitigar el desequilibrio de clases podrían mejorar el rendimiento general del sistema, convirtiéndolo en una herramienta más fiable para la monitorización continua de actividades en entornos domésticos.

5. Conclusiones y trabajos futuros

Este estudio presenta un sistema innovador para la clasificación continua de actividades diarias en pacientes con temblor, utilizando SVMs entrenadas con datos segmentados en diferentes tamaños de ventana. Con una precisión global del 73%, el sistema muestra potencial para clasificar tareas en el entorno doméstico, abarcando secuencias de 5 a 70 minutos, con el objetivo de extender esta capacidad para analizar señales continuas. Aunque el sistema es prometedor, aún enfrenta desafíos relacionados con el desequilibrio de clases y la identificación precisa de tareas bimanuales. Trabajos futuros deberán incluir pruebas con pacientes con EP y TE en sus hogares, así como la recolección de más datos. Esto permitirá explorar técnicas de aprendizaje profundo, como redes recurrentes, que se proponen como una línea de investigación futura. Estas mejoras podrían incrementar significativamente el rendimiento y la precisión del sistema a medida que el conjunto de datos se expanda.

Agradecimientos

Este trabajo ha sido parcialmente financiado por el Programa Nacional de Investigación Científica, Técnica y de Innovación 2021-2023 en el contexto del proyecto NETremor (TED2021-130174B-C32) y la Universidad Politécnica de Madrid en el marco del proyecto Técnicas de Control y Análisis de Variables Cinemáticas Aplicadas a la Extracción de Patrones de Movimiento (RP2109560107).

Referencias

- [1] Anouti A, Koller WC, City K. Articles Tremor Disorders Diagnosis and Management. 1995 jun.
- [2] Thanvi B, Lo N, Robinson T. Essential tremor - The most common movement disorder in older people. Vol. 35, Age and Ageing. 2006. p. 344-9.
- [3] Demrozi F, Bacchin R, Tamburin S, Cristani M, Pravadelli G. Toward a Wearable System for Predicting Freezing of Gait in People Affected by Parkinson's Disease. IEEE J Biomed Health Inform. septiembre de 2020;24(9):2444-51.
- [4] Serrano JI, Lambrecht S, del Castillo MD, Romero JP, Benito-León J, Rocon E. Identification of activities of daily living in tremorous patients using inertial sensors. Expert Syst Appl. 15 de octubre de 2017;83:40-8.
- [5] Gironell A, Marín-Lahoz J, Póveda S. Essential tremor: Update of therapeutic strategies. Medicina Clínica (English Edition). junio de 2024;162(12):599-605.
- [6] Iluz T, Gazit E, Herman T, Sprecher E, Brozgol M, Giladi N, et al. Automated detection of missteps during community ambulation in patients with Parkinson's disease: A new approach for quantifying fall risk in the community setting. J Neuroeng Rehabil. 3 de abril de 2014;11(1).
- [7] Vanmechelen I, Haberehner H, De Vleeschhauwer J, Van Wouterghem E, Feys H, Desloovere K, et al. Assessment of movement disorders using wearable sensors during upper limb tasks: A scoping review. Vol. 9, Frontiers in Robotics and AI. Frontiers Media S.A.; 2023.
- [8] Sigcha L, Borzi L, Amato F, Rechichi I, Ramos-Romero C, Cárdenas A, et al. Deep learning and wearable sensors for the diagnosis and monitoring of Parkinson's disease: A systematic review. Expert Syst Appl. noviembre de 2023; 229:120541.
- [9] Jiang B, Han JJ, Kim J. A Wearable In-home Tremor Assessment System via Virtual Reality Environment for the Activities in Daily Lives (ADLs). En: 2022 44th Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC) [Internet]. IEEE; 2022 [citado 25 de marzo de 2024]. p. 1117-20. Disponible en: <https://ieeexplore.ieee.org/document/9871008>
- [10] Banos O, Galvez JM, Damas M, Pomares H, Rojas I. Window Size Impact in Human Activity Recognition. Sensors. 9 de abril de 2014;14(4):6474-99.
- [11] Yala N, Fergani B, Fleury A. Towards improving feature extraction and classification for activity recognition on streaming data. J Ambient Intell Humaniz Comput. 24 de abril de 2017;8(2):177-89.
- [12] Nurwulan N, Jiang B. Window Selection Impact in Human Activity Recognition. International Journal of Innovative Technology and Interdisciplinary Sciences. 28 de febrero de 2020;3:381-94.
- [13] Ruiz-Vitte A, Carbone E, Larraga B, Rocon E, Gutiérrez Á. The importance of integral time length windows for the classification of activities of daily living based on Machine Learning techniques. En: Proceedings of the XLI Congreso Anual de la Sociedad Española de Ingeniería Biomédica. Cartagena; 2023.
- [14] Gallego J, Rocon E, Roa JO, Moreno J, Pons JL. Real-Time Estimation of Pathological Tremor Parameters from Gyroscope Data. Sensors. 16 de marzo de 2010;10(3):2129-49.
- [15] Google Fitbit. Advanced Health Smartwatch | Fitbit Sense.
- [16] Martín Mg. Universidad Politécnica De Madrid Escuela Técnica Superior De Ingenieros De Telecomunicación Contributions To Human Motion Modeling And Recognition Using Non-Intrusive Wearable Sensors Tesis Doctoral. 2022.
- [17] Pedregosa F, Varoquaux G, Gramfort A, Michel V, Thirion B, Grisel O, et al. Scikit-learn: Machine Learning in Python. Journal of Machine Learning Research. 2011;12:2825-30
- [18] Noble WS. What is a support vector machine? Nat Biotechnol. diciembre de 2006;24(12):1565-7