

# Predicción de la enfermedad de Parkinson utilizando redes neuronales convolucionales

Renata S. Guatelli, Verónica I. Aubin, Silvia N. Pérez

Departamento de Informática e Investigaciones Tecnológicas  
Universidad Nacional de La Matanza

rguatelli@unlam.edu.ar; vaubin@unlam.edu.ar; sperez@unlam.edu.ar

## RESUMEN

La enfermedad de Parkinson (EP) es un desorden neurodegenerativo del sistema nervioso, de causa desconocida y curso crónico, progresivo e irreversible. En la actualidad se asume que los cambios patofisiológicos que permiten apreciar los síntomas de la enfermedad, no son visibles hasta al menos cuatro años luego de su inicio. Por esta causa, se buscan métodos alternativos que permitan detectar la enfermedad en forma temprana. Dado que las deficiencias del habla es uno de los síntomas de la enfermedad, esto puede dar origen a un biomarcador para el diagnóstico temprano y el monitoreo de la enfermedad.

Este trabajo propone un estudio a partir del aprendizaje profundo de los espectrogramas obtenidos de señales de voz grabadas con celulares. Como objetivo se plantea realizar aportes al diagnóstico de EP, contribuyendo asimismo al conocimiento de las características de la voz afectadas por la enfermedad.

Para tal fin se creará una base de datos de espectrogramas de los segmentos de audio que mejor permitan caracterizar la voz de los EP. Se desarrollarán modelos de redes neuronales convolucionales con distintas arquitecturas para distinguir los EP de los pacientes sanos, utilizando la validación adecuada para las características de dichos datos.

**Palabras clave:** *espectrograma, red neuronal, enfermedad de Parkinson.*

## CONTEXTO

Esta investigación se realiza como parte del Proyecto de investigación PROINCE-C224, de la Universidad Nacional de La Matanza. Se desarrolla conjuntamente entre el

Departamento de Salud y el de Ingeniería e Investigaciones Tecnológicas de la UNLaM.

Propone el análisis estadístico de señales acústicas de la voz como método objetivo y no invasivo para el diagnóstico precoz de la enfermedad de Parkinson (EP).

El mismo es continuación del proyecto PROINCE-C199 de la Universidad Nacional de La Matanza, llevado a cabo durante los años 2017 y 2018.

## 1. INTRODUCCIÓN

La enfermedad de Parkinson idiopática (EPI) es un desorden neurodegenerativo del sistema nervioso, de causa desconocida y curso crónico, progresivo e irreversible. Las manifestaciones motoras marcadoras de la enfermedad están representadas por temblor de reposo, rigidez y bradicinesia [1]. En su evolución natural se agregan otros síntomas como: trastornos de la marcha, desequilibrio y alteraciones de la voz lo que implica un marcado impacto en la calidad vida limitando la movilidad con caídas frecuentes y aislamiento social como resultado de severa disartria [2].

Su proceso de diagnóstico y seguimiento, aún hoy en día, es lento y complicado al no existir ninguna prueba específica. En su defecto, el personal sanitario tiene que evaluar concienzudamente el historial clínico, síntomas y exámenes físicos y neurológicos de cada paciente, y ofrecer un diagnóstico a menudo poco preciso. Es por ello, que se estima que hasta un 25% de los diagnosticados como enfermos de Parkinson realmente padecen otra enfermedad, es decir, el proceso actual provoca numerosos falsos positivos.

La rigidez asociada afecta al sistema respiratorio encargado de facilitar el flujo y la presión de aire necesarios para la emisión

vocal lo que se traduce en una disminución de presión sub-glótica. Esto se refleja en un tono inadecuadamente bajo, resonancia gutural posterior, intensidad baja, voz monótona e hipo nasalidad [3]. Las deficiencias del habla son comúnmente uno de los síntomas tempranos de la enfermedad, por lo que puede ser un buen biomarcador para el apoyo diagnóstico y el monitoreo de la enfermedad. Se estima que entre un 60-80% de paciente con EPI presentan alteraciones de la voz caracterizados por cambios en frecuencia, duración e intensidad [4], [5]. Estas alteraciones se confunden en muchas ocasiones con los cambios naturales de los adultos mayores en relación con la presbifonía [6], [7] o estados depresivos [8]. Los pacientes con EPI reflejan la disminución en la capacidad de los músculos laríngeos para mantener una posición fija en la pronunciación sostenida de las vocales [9]. También se observa una disminución en la capacidad para producir prosodia.

Entre los hallazgos más frecuentes en la valoración de la voz de pacientes con EPI se encuentra el incremento u otra variación de la frecuencia fundamental (F0), la reducción del tiempo de producción de vocales, el aumento del VOT (del inglés Voice Onset Time, tiempo de inicio de la sonoridad de la vocal que sigue a la consonante), la disminución de la intensidad en la fonación, así como el decrecimiento del MPT (tiempo máximo de fonación), perturbación del tono Jitter e intensidad Shimmer, la razón ruido/armónicos, también ha mostrado diferencias significativas [10]. Metter y Hanson pusieron de manifiesto que el incremento en la F0 es paralelo a la gravedad de los síntomas y al avance de la enfermedad [11].

La alteración en prosodia expresiva está documentada y justificada por la reducción [12], variabilidad [11] e intensidad [13], [14], de la F0 en tareas de lectura de párrafos en los que los sujetos debían imitar frases acentuando su contenido emocional. Estos resultados ayudan en la búsqueda de los efectos que provoca esta enfermedad en la respiración, fonación, articulación y prosodia [15].

En los últimos años se han propuesto varios sistemas computarizados para identificar los síntomas tempranos de la enfermedad de Parkinson, monitoreando los patrones de voz [16], [17], [18] y [19]. En la Universidad de Oxford, [20], [18], definieron un conjunto de medidas de disfonía y compararon los resultados de usar cuatro algoritmos de selección de características y clasificación binaria. Los integrantes del grupo de investigación Informática Aplicada al Procesado de Señal de la Universidad Politécnica de Madrid (UPM) en colaboración con el Centro de Tecnología Biomédica, han desarrollado un simulador biomecánico para modelar las particularidades del aparato fonador y han evaluado diferentes propuestas de extracción de parámetros para detectar la existencia de patologías utilizando la señal de voz de los pacientes [21].

A través del análisis articulatorio de la repetición rápida de sílabas como /pa-ta-ka/, calculando características temporales y espectrales extraídas en los segmentos de tiempo de inicio de voz [22], propone diseñar un sistema experto para la detección temprana de la EP.

Las redes neuronales convolucionales (CNN) han permitido mejoras significativas en aplicaciones en visión por computadora, tales como detección de objetos, reconocimiento de rostros, clasificación de imágenes, procesamiento del lenguaje natural, procesamiento de audio, entre otros [23]. Los algoritmos de aprendizaje profundo o deep learning buscan en el aprendizaje automático características de alto de alto nivel a partir de grandes volúmenes de datos [24].

Los avances en las TICs (Tecnologías de la Información y Comunicación), han abierto la posibilidad del seguimiento online de la enfermedad mediante el análisis de grabaciones de voz recogidas con teléfonos móviles. Esto permite disponer de un gran número de datos de pacientes con mínimo trabajo de recopilación. Estas tecnologías junto a las aplicaciones de aprendizaje profundo hacen viable la asistencia sanitaria móvil.

Una señal de audio se puede representar visualmente como un espectrograma que captura el contenido de frecuencia variable de la señal, ver Figura 1. Es una representación en tres dimensiones, tiempo, frecuencia y amplitud de la distribución de energía de una señal. En la imagen se visualiza la intensidad del sonido mediante las variaciones de color, a lo largo del tiempo (eje horizontal) en función de la frecuencia (eje vertical).

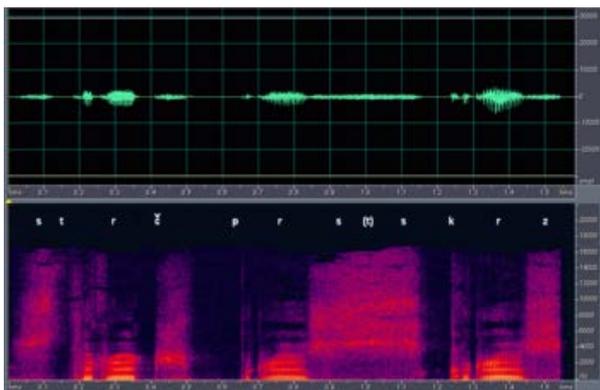


Figura 1. Espectrograma generado a partir de una señal de voz.

En este trabajo de investigación se propone discernir entre pacientes con EP y personas sanas a partir del uso de una CNN aplicada a los espectrogramas de señales de voz obtenidas con teléfonos móviles. Los sonidos que se utilizarán corresponden a los de la base de datos mPower, que contiene más de 64.000 grabaciones de voz, capturadas con teléfonos celulares discriminados entre enfermos de Parkinson y controles sanos. También se realizarán pruebas sobre la base de datos generada en el proyecto C-224, grabada en un ambiente controlado, siguiendo el protocolo de grabación acordado con profesionales del Hospital Nacional Alejandro Posadas, del Hospital Bernardino Rivadavia y con el asesoramiento del Dr. Jorge Gurlekian.

## 2. LINEA DE INVESTIGACION

La línea de investigación que se desarrolla pretende detectar EP utilizando un método no invasivo. En la literatura, existe un consenso sobre que el análisis acústico resulta útil para tal fin.

La propuesta aquí presentada consiste en obtener modelos que permitan colaborar con el

diagnóstico temprano de la enfermedad, estudiando la aplicación del aprendizaje profundo de los espectrogramas, obtenidos a partir de registros de señales de voz. Dichas señales corresponden a grabaciones realizadas a través de teléfonos celulares, ya que esta tecnología promueve la asistencia sanitaria móvil.

Asimismo, se explorarán distintas particularidades de la voz de los EP, pre-procesando los audios para extraer distintos fragmentos (por ej. comienzo, final, etc.) que permitan evaluar a través de las CNN con cuál de ellos se logra una mejor clasificación.

## 3. RESULTADOS

Como resultado de esta investigación se obtendrán modelos de detección de la enfermedad de Parkinson, utilizando espectrogramas obtenidos a partir de datos de voz recogidos a través de grabaciones de teléfonos móviles, implementando algoritmos de redes neuronales basadas en Deep Learning. La metodología utilizada contempla tres etapas:

- Pre-procesado de los audios y creación de la base de datos
- Construcción de la red neuronal convolucional
- Procesamiento de la Base de Datos

**En la primera etapa** se procesarán los audios. Se seleccionarán y aplicarán distintos algoritmos para eliminar el ruido de fondo. Se seleccionarán los segmentos de audio significativos para el estudio del reconocimiento de EP y se crearán los espectrogramas de los audios seleccionados.

**En la segunda etapa**, una vez tratados los datos iniciales para introducirlos en la red neuronal convolucional, se desarrollarán los modelos correspondientes a distintas arquitecturas propuestas en la literatura.

Por último, en la **tercera etapa**, se evaluarán y compararán los modelos generados. Para la evaluación de cada modelo se seleccionará el tipo de validación más adecuada teniendo en cuenta las características de los datos, asegurándose que el conjunto de entrenamiento y de test sean representativos por sí mismos. Para comparar los modelos

desarrollados se implementarán diferentes indicadores de desempeño como matriz de confusión y curva ROC, entre otros.

El proyecto de investigación se desarrolla conjuntamente entre el Departamento de Salud y el de Ingeniería e Investigaciones Tecnológicas de la UNLaM. Los resultados obtenidos impactarán sobre los recursos disponibles de ambos departamentos. Estos implican un avance del conocimiento en el campo de la detección de enfermedades neurodegenerativa, posibilitando la formación de recursos humanos capacitados en una nueva línea de investigación para esta casa de altos estudios.

Se propone además realizar vínculos con grupos de investigación de Argentina o el exterior, que se encuentren trabajando en temas afines a este proyecto.

Se divulgarán los resultados obtenidos y se hará transferencia de tecnología a potenciales usuarios del área de salud. Se prevé como primeros beneficiarios a la comunidad de La Matanza y luego a la comunidad en general, ya que se busca un método de bajo costo para el diagnóstico precoz del Parkinson.

#### 4. FORMACIÓN DE RECURSOS HUMANOS

En el proyecto participan estudiantes de grado y jóvenes graduados de la UNLaM, quienes han obtenido becas de iniciación a la investigación.

En base a esta investigación está previsto el desarrollo de una Tesis de Maestría en Informática a cargo de Lic. Renata Silvia Guatelli.

#### 5. BIBLIOGRAFÍA

[1] Jankovic, J. (2008). Parkinson's disease: clinical features and diagnosis. *Journal of neurology, neurosurgery & psychiatry*, 79(4), 368-376.

[2] Köllensperger, M., Geser, F., Seppi, K., Stampfer-Kountchev, M., Sawires, M., Scherfler, C., & Pallecchia, M. T. (2008). Red flags for multiple system atrophy. *Movement disorders: official journal of the Movement Disorder Society*, 23(8), 1093-1099.

official journal of the Movement Disorder Society, 23(8), 1093-1099.

[3] Landázuri, E., Villamil, L., & Delgado, L. (2007). Parámetros acústicos de la voz en personas con enfermedad de Parkinson. *Umbral Científico*, (11), 90-103.

[4] De Letter, M., Santens, P., & Van Borsel, J. (2005). The effects of levodopa on word intelligibility in Parkinson's disease. *Journal of communication disorders*, 38(3), 187-196.

[5] Törnqvist, A. L., Schalén, L., & Rehnström, S. (2005). Effects of different electrical parameter settings on the intelligibility of speech in patients with Parkinson's disease treated with subthalamic deep brain stimulation. *Movement disorders*, 20(4), 416-423.

[6] Fernández González, S., Ruba San Miguel, D., Marques Girbau, M., Sarrasqueta Sáenz, L. (2006). "Voz del anciano" *Revista de Medicina de la Universidad de Navarra*, vol. 50, Nº 3, 44-48.

[7] Morrison, M., & Ramage, L. (1996). Tratamiento de los trastornos de la voz. Elsevier España.

[8] Cummings, J. L., & Masterman, D. L. (1999). Depression in patients with Parkinson's disease. *International journal of geriatric psychiatry*, 14(9), 711-718.

[9] Gallena, S., Smith, P. J., Zeffiro, T., & Ludlow, C. L. (2001). Effects of levodopa on laryngeal muscle activity for voice onset and offset in Parkinson disease. *Journal of Speech, Language, and Hearing Research*.

[10] Jiménez-Jiménez, F. J., Gamboa, J., Nieto, A., Guerrero, J., Orti-Pareja, M., Molina, J. A., & Cobeta, I. (1997). Acoustic voice analysis in untreated patients with Parkinson's disease. *Parkinsonism & Related Disorders*, 3(2), 111-116.

[11] Metter, E. J., & Hanson, W. R. (1986). Clinical and acoustical variability in hypokinetic dysarthria. *Journal of communication disorders*, 19(5), 347-366.

[12] Flint, A. J., Black, S. E., Campbell-Taylor, I., Gailey, G. F., & Levinton, C. (1992). Acoustic analysis in the differentiation of Parkinson's

disease and major depression. *Journal of Psycholinguistic Research*, 21(5), 383-399.

[13] Hertrich, I., & Ackermann, H. (1995). Gender-specific vocal dysfunctions in Parkinson's disease: electroglottographic and acoustic analyses. *Annals of Otology, Rhinology & Laryngology*, 104(3), 197-202.

[14] J. Holmes, R., M. Oates, J., J. Phyland, D., & J. Hughes, A. (2000). Voice characteristics in the progression of Parkinson's disease. *International Journal of Language & Communication Disorders*, 35(3), 407-418.

[15] Aguilera Pacheco, O. R., Escobedo Beceiro, D. I., Sanabria Macías, F., & Nuñez Lahera, I. (2015). Alteración de parámetros acústicos de la voz y el habla en la enfermedad de Parkinson. In XIV Simposio Internacional de Comunicación Social. *Comunicación Social: Retos y Perspectivas* (Vol. 2).

[16] Asgari, M., & Shafran, I. (2010). Extracting cues from speech for predicting severity of parkinson's disease. In *2010 IEEE International Workshop on Machine Learning for Signal Processing* (pp. 462-467). IEEE.

[17] Sakar, C.O and Kursun, O. (2009) 'Telediagnosis of Parkinson's disease using measurements of dysphonia', *Journal of Medical Systems*, DOI: 10.1007/s10916-009-9272-y.

[18] Tsanas, A., Little, M. A., McSharry, P. E., Spielman, J., & Ramig, L. O. (2012). Novel speech signal processing algorithms for high-accuracy classification of Parkinson's disease. *IEEE transactions on biomedical engineering*, 59(5), 1264-1271.

[19] Alhussein, M., & Muhammad, G. (2018). Voice pathology detection using deep learning on mobile healthcare framework. *IEEE Access*, 6, 41034-41041.

[20] Tsanas, A., Little, M. A., McSharry, P. E., & Ramig, L. O. (2010). Nonlinear speech analysis algorithms mapped to a standard metric achieve clinically useful quantification of average Parkinson's disease symptom severity. *Journal of the royal society interface*, 8(59), 842-855

[21] Díaz-Pérez, F., García-Nieto, E., Ros, A., & Claramunt, R. (2014). Best estimation of spectrum

profiles for diagnosing femoral prostheses loosening. *Medical engineering & physics*, 36(2), 233-238.

[22] Montaña, D., Campos-Roca, Y., & Pérez, C. J. (2018). A diadochokinesis-based expert system considering articulatory features of plosive consonants for early detection of Parkinson's disease. *Computer methods and programs in biomedicine*, 154, 89-97.

[23] Deng, L., & Yu, D. (2014). Deep learning: methods and applications. *Foundations and Trends® in Signal Processing*, 7(3-4), 197-387.

[24] Tan, C., Sun, F., Kong, T., Zhang, W., Yang, C., & Liu, C. (2018, October). A survey on deep transfer learning. In *International conference on artificial neural networks* (pp. 270-279). Springer, Cham.