

**DETECCIÓN DE PATRONES ANORMALES ASOCIADOS A LA  
ENFERMEDAD DEL PARKINSON EN SECUENCIAS DE MOVIMIENTO  
OCULAR USANDO REDES GENERATIVAS ADVERSARIAS**

**JUAN DAVID NIÑO RODRIGUEZ**

**UNIVERSIDAD INDUSTRIAL DE SANTANDER  
FACULTAD DE INGENIERÍAS FÍSICOMECÁNICAS  
ESCUELA DE INGENIERÍA DE SISTEMAS E INFORMÁTICA  
BUCARAMANGA**

**2021**

**DETECCIÓN DE PATRONES ANORMALES ASOCIADOS A LA  
ENFERMEDAD DEL PARKINSON EN SECUENCIAS DE MOVIMIENTO  
OCULAR USANDO REDES GENERATIVAS ADVERSARIAS**

**JUAN DAVID NIÑO RODRIGUEZ**

Una tesis presentada en cumplimiento de los requisitos para el grado de:  
**Ingeniero de Sistemas e Informática**

**Director:**

**Fabio Martínez Carrillo**

**Ph.D en Ingeniería de Sistemas y Computación**

**UNIVERSIDAD INDUSTRIAL DE SANTANDER  
FACULTAD DE INGENIERÍAS FÍSICOMECAÑICAS  
ESCUELA DE INGENIERÍA DE SISTEMAS E INFORMÁTICA  
BUCARAMANGA**

**2021**

## AGRADECIMIENTOS

El autor expresa su agradecimiento:

Primero agradezco a mi madre Gladys Rodriguez, por haber sido un pilar fundamental durante toda mi vida y por su apoyo en todos los momentos de gran dificultad, y a mi padre Manuel Niño, por haber sido un ejemplo a seguir en mi vida y por siempre creer en mí durante mi etapa de estudiante.

A mis hermanos Diego y Sergio por su gran apoyo durante mi carrera universitaria y por estar siempre presentes cuando más los necesito.

A mi profesor y director de proyecto de grado, Fabio Martínez, por ser la más grande inspiración durante mi carrera universitaria, y por impulsarme siempre a alcanzar metas más altas. Por mostrarme el camino a tomar en la academia y la investigación, y por su apoyo y paciencia ilimitada.

Al ingeniero Santiago Gomez, que ha sido tutor durante mi proyecto de grado, por ser una persona confiable, afable y respetuosa, y por ser una gran figura de ejemplo para esta etapa final de mi carrera, cuyos consejos siempre fueron importantes para seguir adelante.

A mis amigos Juan Pablo, Santiago, Jorge, Manuel y María Fernanda por ser seres humanos invaluable, y por haber acompañado gran parte de mi vida, durante el colegio y durante mi vida universitaria. Mi logro más grande en la vida fue haberlos conocido, y haber podido gozar de momentos de alegría y tristeza junto a ellos.

Y por último a mis compañeros del grupo de investigación Bivl2ab, cuyos consejos han sido escalones fundamentales en esta investigación.

## CONTENIDO

|                                                                               | Pág       |
|-------------------------------------------------------------------------------|-----------|
| <b>INTRODUCCIÓN</b> . . . . .                                                 | <b>10</b> |
| <b>1. MARCO TEÓRICO Y TRABAJOS PREVIOS</b> . . . . .                          | <b>13</b> |
| 1.1. PARKINSON Y LOS TEMBLORES OCULARES . . . . .                             | 13        |
| 1.2. MÉTODOS DE CUANTIFICACIÓN PARKINSONIANOS . . . . .                       | 15        |
| 1.3. REDES GENERATIVAS ADVERSARIAS . . . . .                                  | 18        |
| 1.3.1. GAN CONDICIONAL . . . . .                                              | 22        |
| 1.4. DETECCIÓN DE ANOMALÍAS . . . . .                                         | 23        |
| 1.4.1. DETECCIÓN DE ANOMALÍAS CON RPCA . . . . .                              | 25        |
| 1.4.2. DETECCIÓN DE ANOMALÍAS CON AUTOENCODERS . . . . .                      | 26        |
| 1.4.3. DETECCIÓN DE ANOMALÍAS CON GANS . . . . .                              | 27        |
| <b>2. PLANTEAMIENTO Y JUSTIFICACIÓN DEL PROBLEMA</b> . . . . .                | <b>30</b> |
| <b>3. OBJETIVOS</b> . . . . .                                                 | <b>31</b> |
| <b>4. ENFOQUE PROPUESTO</b> . . . . .                                         | <b>32</b> |
| 4.1. SEGMENTACIÓN DE VIDEOS . . . . .                                         | 32        |
| 4.2. MAGNIFICACIÓN DE ACELERACIÓN DE VIDEO . . . . .                          | 33        |
| 4.3. ARQUITECTURA GANOMALY PARA REPRESENTAR PATRONES PARKINSONIANOS . . . . . | 36        |
| 4.3.1. SUBRED GENERADORA . . . . .                                            | 37        |
| 4.3.2. SUBRED CODIFICADORA . . . . .                                          | 38        |
| 4.3.3. SUBRED DISCRIMINADORA . . . . .                                        | 38        |
| 4.4. ENTRENAMIENTO DEL MODELO GENERATIVO GANOMALY . . . . .                   | 39        |

|                                                 |           |
|-------------------------------------------------|-----------|
| <b>5. DISEÑO EXPERIMENTAL</b> . . . . .         | <b>41</b> |
| 5.1. CONJUNTO DE DATOS . . . . .                | 41        |
| 5.2. CONFIGURACIÓN DE LA ESTRATEGIA . . . . .   | 41        |
| 5.3. MÉTRICAS DE VALIDACIÓN . . . . .           | 42        |
| <b>6. EVALUACIÓN Y RESULTADOS</b> . . . . .     | <b>45</b> |
| <b>7. CONCLUSIONES Y PERSPECTIVAS</b> . . . . . | <b>50</b> |
| <b>BIBLIOGRAFÍA</b> . . . . .                   | <b>52</b> |

## TABLA DE FIGURAS

|                                                            | Pág |
|------------------------------------------------------------|-----|
| Figura 1. Ejemplos de movimientos . . . . .                | 17  |
| Figura 2. Red generativa adversaria . . . . .              | 19  |
| Figura 3. Entrenamiento del modelo discriminador . . . . . | 20  |
| Figura 4. Entrenamiento del modelo generativo . . . . .    | 21  |
| Figura 5. Arquitectura de la CGAN . . . . .                | 23  |
| Figura 6. Idea general de identificación anómala . . . . . | 32  |
| Figura 7. Segmentación del video . . . . .                 | 33  |
| Figura 8. Arquitectura de Ganomaly . . . . .               | 37  |
| Figura 9. Subred generadora . . . . .                      | 38  |
| Figura 10. Subred codificadora . . . . .                   | 39  |
| Figura 11. Subred discriminadora . . . . .                 | 39  |
| Figura 12. Curva ROC promedio del experimento . . . . .    | 47  |
| Figura 13. Boxplot del experimento . . . . .               | 48  |
| Figura 14. Matriz de confusión del experimento . . . . .   | 48  |
| Figura 15. Substracción de imágenes . . . . .              | 49  |

## TABLA DE CUADROS

|                                                                                   | <b>Pág</b> |
|-----------------------------------------------------------------------------------|------------|
| Tabla 1. Resultados de los tres experimentos representados en promedios . . . . . | 45         |
| Tabla 2. Resultados del análisis estadístico usando ANOVA . . . . .               | 46         |

## RESUMEN

**TÍTULO:** DETECCIÓN DE PATRONES ANORMALES ASOCIADOS A LA ENFERMEDAD DEL PARKINSON EN SECUENCIAS DE MOVIMIENTO OCULAR USANDO REDES GENERATIVAS ADVERSARIAS \*

**AUTOR:** JUAN DAVID NIÑO RODRIGUEZ \*\*

**PALABRAS CLAVE:** Enfermedad del Parkinson, Fijación Ocular, Redes Generativas Adversarias, Detección de anomalías

**DESCRIPCIÓN:** Las alteraciones oculomotoras constituyen un biomarcador prometedor para la detección y caracterización de la enfermedad del Parkinson (EP), incluso en etapas prodrómicas. Sin embargo, actualmente solo las trayectorias globales y simplificadas de movimiento ocular, obtenidas de dispositivos de rastreo, son usadas para aproximar la compleja cinemática de la función oculomotora. Además, la adquisición de tales señales requiere de calibración sofisticada y ajustes intrusivos. La cuantificación de patrones oculares en secuencias de video ha sido una alternativa para complementar este análisis, involucrando además herramientas de aprendizaje de máquina para clasificaciones de patrones patológicos. Estos enfoques, sin embargo, se basan principalmente en modelos discriminatorios que requieren condiciones estrictas en cuanto al número y balance de los datos de entrenamiento. Este trabajo construye un novedoso descriptor de video para la cuantificación de patrones parkinsonianos, durante tareas de fijación ocular. El descriptor fue desarrollado en un marco de detección de anomalías, donde se asume únicamente la disponibilidad de muestras de una clase de interés. Por lo tanto, el enfoque presentado se enfoca solo en la representación modal del Parkinson, considerando todas las otras clases de muestra como una anomalía de la distribución, y logrando así su clasificación. Este enfoque fue evaluado en total un total de 13 pacientes control y 13 pacientes Parkinson. En la tarea de clasificación, el biomarcador digital propuesto alcanzó una sensibilidad y especificidad de 0.965 y 0.685, respectivamente. Una prueba estadística muestra diferencias significantes ( $p < 0.05$ ) entre las clases ´predichas, codificadas por el descriptor propuesto, evidenciando una discriminación entre pacientes Parkinson y control.

---

\* Trabajo de investigación

\*\* Facultad de Ingenierías Físico-Mecánicas. Escuela de Ingeniería de Sistemas e Informática. Director: Fabio Martínez Carrillo, Ph.D.

## ABSTRACT

**TITLE:** ANOMALY DETECTION OVER PARKINSON-AFFECTED EYE SLICES USING GENERATIVE ADVERSARIAL NETWORKS. \*

**AUTHOR:** JUAN DAVID NIÑO RODRIGUEZ \*\*

**KEYWORDS:** Parkinsons disease, Ocular Fixation, Generative Adversarial Networks, Anomaly detection.

**DESCRIPTION:** Oculomotor alterations constitute a promising biomarker to detect and characterize Parkinson's disease, even in prodromal stages. However, currently only global and simplified eye movement trajectories, obtained from tracking devices, are used to approximate the complex and hidden kinematic relationships of the oculomotor function. Besides, the acquisition of such signals require of sophisticated calibration and intrusive setups. The quantification of ocular patterns in video sequences have risen as an alternative to complement this analysis, also involving machine learning tools for the classification of pathological patterns. These schemes, however, are primarily based on discriminatory models that require strict conditions regarding the number and balancing of the training data. This work constructs a novel video descriptor for the quantification of parkinsonian patterns, during ocular fixation tasks. The descriptor was developed in a anomaly detection framework, where only the availability of samples from one class of interest is assumed. Therefore, the proposed approach focuses only on Parkinson's mode representation, considering any other class sample as an anomaly of the distribution, thus obtaining its classification. This approach was evaluated for an ocular fixation task, in a total of 13 control subjects and 13 patients on different stages of the disease. In a classification task, the proposed digital biomarker achieved an average sensibility and specificity of 0.965 and 0.625, respectively. A statistical test shows significant differences ( $p < 0.05$ ) among predicted classes, coded from the proposed biomarker, evidencing a discrimination between patients and control subjects.

---

\* Research work

\*\* Faculty of Physical-Mechanical Engineering. School of Systems and Computer Engineering. Advisor: Fabio Martínez Carrillo

## INTRODUCCIÓN

La enfermedad del Parkinson (EP) es el segundo desorden neurodegenerativo más común en el mundo. La incidencia reportada indica una afectación típica entre el 2 y el 3% de la población mayor a 65 años de edad, con una expectativa de duplicarse para el 2030 <sup>1</sup>. Los síntomas de la EP son primariamente motores, involucrando alteraciones progresivas e involuntarias sobre la movilidad de diferentes segmentos corporales, como el temblor, la bradicinesia (movimiento lento) y la rigidez. Estos impedimentos físicos son la consecuencia directa de un declive gradual en niveles de dopamina, una biomolécula que el cerebro usa para la neurotransmisión de comandos motores.<sup>2</sup>

A pesar de avances recientes en neurología y genética, la pérdida de células de secreción de dopamina continúa siendo mal entendido, resultando en procedimientos de diagnóstico que únicamente se fían de alteraciones motoras notables. En la práctica clínica, tales procedimientos llevan a cabo pruebas observacionales específicas, como la Escala Unificada de la Enfermedad del Parkinson (UPDRS) y la escala Hoehn-Yahr, que correlaciona el comportamiento motriz del paciente con diferentes etapas de la enfermedad. <sup>3</sup> Sin embargo, esta evaluación es propensa a cierta parcialidad y subjetividad debido a la alta variabilidad interpersonal del movimiento del paciente y la experiencia y percepción particular del médico.<sup>4</sup> En la actualidad, los modelos cuantitativos y reproducibles, se presentan como una alternativa viable para reducir estos

---

<sup>1</sup> POEWE, Werner. “Parkinson’s disease”. En: *Nature reviews Disease primers 3* (2017), pág. 17013.

<sup>2</sup> JANKOVIC, Joseph. “Parkinson’s disease: clinical features and diagnosis”. En: *Journal of neurology, neurosurgery and psychiatry* 79.4 (2008), págs. 368-376.

<sup>3</sup> VENUTO, Charles. “A review of disease progression models of Parkinson’s disease and applications in clinical trials”. En: *Movement Disorders* 31.4 (2016), págs. 368-376.

<sup>4</sup> COPETTI, Massimiliano, RIZZO, Giovanni y ARCUTI, Simona. “Accuracy of clinical diagnosis of Parkinson disease A systematic review and meta-analysis”. En: *Neurology* 86.6 (2016), págs. 566-576.

problemas y obtener una evaluación objetiva de la EP. Además, estos modelos podrían llegar a tener un grado de precisión/sensibilidad considerable en las etapas tempranas de la enfermedad, donde el grado de incapacidad motora es muy sutil, lo cual deriva en un error de diagnóstico cuantificado en más del 24 %, incluso en centros especializados.<sup>1</sup> Recientemente, en la literatura se han reportado importantes avances en cuanto a la detección temprana de la enfermedad, partiendo de la cuantificación de patrones de movimiento ocular. Estos movimientos pueden estar asociados a deficiencias en el control propio del ojo <sup>5</sup> o ser la consecuencia de movimientos que se reflejan por la inestabilidad del tronco y la cabeza. <sup>6</sup> Estos análisis sin embargo hoy en día requieren complejos instrumentos y protocolos de captura para la cuantificación de correlaciones motoras, asociadas con la enfermedad .

El aprendizaje de máquina (AM) y los métodos de reconocimiento de patrones han permitido mejorar el diagnóstico y monitoreo del paradigma de la EP por el modelamiento de diversos síntomas cinemáticos.<sup>7</sup> Específicamente, los trabajos relacionados han reportado discriminación significativa entre la EP y patrones capturados en sujetos control, aplicando algoritmos de AM sobre señales de sensores sobre cuerpo.<sup>8 9</sup> Otros enfoques de buen rendimiento han considerado

---

<sup>5</sup> WETZEL, Paul GITCHEL, George y BARON, Mark. “Pervasive ocular tremor in patients with Parkinson disease”. En: *Archives of neurology* 69.8 (2012), págs. 1011-1017.

<sup>6</sup> BUCKWELL, D. KASKI, D. SAIFEE, T.A.. “Ocular tremor in Parkinson’s Disease Is due to head oscillation”. En: *Movement Disorders* 28 (2013).

<sup>7</sup> BELIC, Minja. “Artificial intelligence for assisting diagnostics and assessment of Parkinson’s disease – A review”. En: *Clinical neurology and neurosurgery* (2019), pág. 105442.

<sup>8</sup> CARAMIA, Carlotta. “IMU-Based Classification of Parkinson’s Disease From Gait: A Sensitivity Analysis on Sensor Location and Feature Selection”. En: *IEEE journal of biomedical and health informatics* 22.6 (2018), págs. 1765-1774.

<sup>9</sup> ABDULHAY, Enas. “Gait and tremor investigation using machine learning techniques for the diagnosis of Parkinson disease”. En: *Future Generation Computer Systems* 83 (2018), págs. 366-373.

el uso de información de video en arquitecturas de aprendizaje profundo.<sup>10 11</sup> Estos esquemas han trabajado principalmente en la clasificación y reconocimiento de marcha en la EP y comportamientos de temblor, enfocándose en extremidades superiores e inferiores. Sin embargo, los síntomas de extremidades motoras aparecen generalmente en etapas medianas y avanzadas de la enfermedad, luego de la pérdida del 50% de las hormonas que secretan dopamina.<sup>12</sup> Con respecto a la cuantificación de patrones oculares, en la literatura se han reportado algunas estrategias para la clasificación y discriminación entre poblaciones diagnosticadas con Parkinson y sujetos control.<sup>13</sup>

Este trabajo introduce un nuevo descriptor de fijación ocular para la caracterización automática de patrones de EP, registrados en video. En este trabajo se resaltan dos aproximaciones metodológicas para el modelamiento de estos patrones. Primero, el análisis basado en anomalías, que pretende hacer representaciones únicamente de la población con EP, siendo cualquier otra muestra discriminada como un valor atípico. Segundo, el uso de esquemas profundos generativos que reportan mayor flexibilidad para aprender patrones ocultos que puedan tener correlación con la enfermedad y eviten el sesgo de las anotaciones referentes a estadios o diagnósticos de los sujetos. En cuanto a la implementación y desarrollo del modelo se validó que un esquema profundo basado en la arquitectura GANomaly permite una representación robusta de la enfermedad. El esquema propuesto además permitió una identificación de los segmentos espacio-temporales que están asociados con la EP. Esta tarea fue evaluada con un conjunto de datos capturados durante la tarea de fijación ocular, en sujetos control y pacientes con la enfermedad del Parkinson.

---

<sup>10</sup> HU, Kun. “Vision-based freezing of gait detection with anatomic patch based representation”. En: *Asian Conference on Computer Vision*. (2018), págs. 564-576.

<sup>11</sup> AJAY, Jerry. “A pervasive and sensor-free deep learning system for Parkinsonian gait analysis”. En: *International Conference on Biomedical and Health Informatics* (2018), págs. 108-111.

<sup>12</sup> SALAT, David. “Challenges of modifying disease progression in prediagnostic Parkinson’s disease”. En: *The Lancet Neurology* 15.6 (2016), págs. 637-648.7.

<sup>13</sup> PERTUZ, S.D. SALAZAR, I y MARTINEZ, F. “Visual amplification of oculomotor signs for parkinsonian patterns recognition”. En: *Computers in biology and medicine* (2020), 8–10.

## 1. MARCO TEÓRICO Y TRABAJOS PREVIOS

### 1.1. PARKINSON Y LOS TEMBLORES OCULARES

Los movimientos oculares han emergido como un posible y prometedor biomarcador de la EP, demostrando una alta correlación con estadios tempranos de la enfermedad. Por ejemplo, *Gorges et al. 2016* encontró que el control ocular del movimiento es disfuncional en la EP, asociado principalmente con deficiencia cortical.<sup>14</sup> *Merel et al. 2017* discutió varios desórdenes y problemas oculares, tales como: superficies oculares secas, diplopía, glaucoma, visión de color y contraste dañado, alucinaciones visuales, entre otras.<sup>15</sup> *Pierpaolo et al. 2018* también encontró varias de estas anormalidades oculares en etapas tempranas de la EP.<sup>16</sup> También, en experimentos especializados se mostró que los pacientes con alteraciones de sueño, que registraban movimientos oculares rápidos tenían una alta probabilidad de desarrollar EP<sup>17</sup>.

Las alteraciones reportadas en la literatura indican una asociación drástica entre la función oculomotora y la neurodegeneración en los procesos claves para desarrollar tareas de locomoción. Una relación que puede ser explicada anatómicamente estudiando las múltiples estructuras

---

<sup>14</sup> MÜLLER, Peter-Hans GORGES, Martin y LULÉ, DOROTHEÉ. “The association between alterations of eye movement control and cerebral intrinsic functional connectivity in Parkinson’s disease”. En: *Brain imaging and behavior* 10.1 (2016), págs. 79-91.

<sup>15</sup> JANSSEN, Sabine EKKER, Merel y SEPPI, Klaus. “Ocular and visual disorders in Parkinson’s disease: Common but frequently overlooked”. En: *Parkinsonism and related disorders* 40 (2017), págs. 1-10.

<sup>16</sup> TURCANO, Pierpaolo. “Early ophthalmologic features of Parkinson’s disease: a review of preceding clinical and diagnostic markers”. En: *Journal of neurology* (2018), págs. 1-9.

<sup>17</sup> SANTAMARIA, Joan, IRANZO, Alex y TOLOSA, Eduardo. “Idiopathic rapid eye movement sleep behaviour disorder: diagnosis, management, and the need for neuroprotective interventions”. En: *The Lancet Neurology* 15.4 (2016), págs. 405-419.

cerebrales y los circuitos neuronales involucrados en el control oculomotor.<sup>18</sup>

Por esta razón, es fundamental detectar y cuantificar las alteraciones oculomotoras, como por ejemplo, la presencia de alteraciones en los patrones de fijación. Estas alteraciones fueron por ejemplo detectadas en 112 pacientes con EP y en 2 de 60 sujetos control, los cuales eventualmente evolucionaron en EP.<sup>5</sup> No obstante, una principal dificultad de la mayoría de las investigaciones para el registro y cuantificación de patrones oculomotores cae en la utilización de protocolos de video-oculografía (VOG)<sup>19</sup>, debido a que su representación se limita a la captura de trayectorias globales y simplificadas de todo el campo de movimiento del ojo. Además estos esquemas de captura son difíciles de calibrar y preparar. Sumado a lo anterior, la instrumentación requiere usualmente contacto con toda el área alrededor de los ojos, afectando el gesto de visión natural. Del mismo modo, alternativas como la eletro-oculografía puede resultar igualmente invasiva en cuanto a la detección y asociación de cambios de potencial relacionados con los movimientos oculares.<sup>13</sup>

Dentro de los patrones oculares, la fijación ocular ha sido un patrón predominante que permite correlacionar alteraciones motoras con la EP, cuando se le solicita a un paciente que mire hacia un punto fijo, durante un intervalo de tiempo específico. Durante este experimento se esperan patrones relativamente bajos de temblor ocular en poblaciones control, mientras que en pacientes con la EP se pueden presentar alteraciones rítmicas y con frecuencias típicas ( $\geq 6Hz$ ), que pueden aumentar en magnitud según los niveles de la profesión de la enfermedad. Con respecto a la fijación ocular, Gitchel et al.<sup>5</sup> reportó la ocurrencia de oscilaciones oculares que poseían las características de frecuencia alrededor de 4 - 7 Hz, similar a esos de los temblores de aquellas personas con EP. Por otra parte, Kaski et. al.<sup>6</sup> mostró que los temblores oculares poseían propiedades espectrales idénticas a aquellos con temblores de cabeza y que el movimiento

---

<sup>18</sup> GOLDBERG, M.E. y WALKER, M.F.. “The control of gaze”. En: *Principles of Neural Science*. (2013), 894–916.

<sup>19</sup> GARCÍA, C.E. LARRAZABAL, A.J. y MARTÍNEZ, F.. “Video-oculography eye tracking towards clinical applications: A review”. En: *Computers in biology and medicine* 108 (2019), 57–66.

ocular ocurría a 180 grados fuera de fase con el del movimiento ocular. En conflicto con este resultado, Gitchel midió los movimientos de ojo y cabeza en 62 de sus pacientes con EP y en 31 pacientes control y encontró que los movimientos de cabeza no contribuían a la inestabilidad ocular en sus resultados. Gitchel también reportó que el temblor ocular persuasivo está presente en pacientes con EP. Este descubrimiento ha generado interés sustancial, y podría servir como un biomarcador para la EP. Por lo tanto, la introducción de nuevas tecnologías para caracterizar tales biomarcadores emergentes son fundamentales para descubrir nuevas relaciones asociadas con la EP.

## 1.2. MÉTODOS DE CUANTIFICACIÓN PARKINSONIANOS

Algunos trabajos en la comunidad de la ciencia de computación han explorado alternativas para la cuantificación de patrones oculares y han propuesto enfoques estadísticos y en base a aprendizaje para predecir el diagnóstico y progreso del Parkinson de manera cuantitativa. Por ejemplo, el trabajo presentado en <sup>20</sup> evaluó las grabaciones sacádicas de 96 pacientes EP, al aplicar clasificadores de random forest para predecir los puntajes individuales de UPRDRS y las dosis de Levodopa (la droga más común para el tratamiento de la EP) de los pacientes. Otro enfoque <sup>21</sup> implementó un modelo binario de regresión logística sobre una combinación de parámetros de fijación y búsqueda sacádica, en el caso de evaluación de 37 pacientes EP y 39 control.

Recientemente, en un estudio compacto de 11 pacientes EP <sup>22</sup>, los algoritmos de árbol de deci-

---

<sup>20</sup> SZYMANSKI, Artur. “Building Intelligent Classifiers for Doctor-Independent Parkinson’s Disease Treatments”. En: *Conference of Information Technologies in Biomedicine* (2016), págs. 267-276.

<sup>21</sup> ZHANG, Yu. “Oculomotor Performances Are Associated With Motor and Non-motor Symptoms in Parkinson’s Disease”. En: *Frontiers in neurology* 9 (2018), pág. 960.

<sup>22</sup> SLEDZIANOWSKI, Albert. “Measurements of Antisaccades Parameters Can Improve the Prediction of Parkinson’s Disease Progression”. En: *Asian Conference on Intelligent Information and Database Systems* (2019), 602–614.

sión son entrenados en métricas antisacádicas para obtener predicciones de los puntajes UPDRS de pacientes . Una versión previa de tal estudio también evaluó las sacadas y los movimientos búsqueda suave usando teoría de conjuntos robusta <sup>23</sup>. Todos los estudios previamente mencionados dependen en esquemas de monitoreo de video-oculografía (VOG) <sup>24</sup>. Estas configuraciones de VOG generalmente emplean cámaras infrarrojas especializadas de alta resolución espacial y temporal, y ofrecen grabaciones confiables y sin ruido del movimiento de los ojos. Sin embargo, el equipo requerido puede ser intrusivo para cubrir toda la región del ojo, es costoso y requiere calibración y procedimientos de configuración. Con el fin de resolver estos problemas, alternativas para la examinación que solo utilizan secuencias de video comunes han sido introducidas recientemente <sup>13 25 26 27</sup>. Tales estrategias modelan las relaciones espacio-temporales de los píxeles basados en diversos algoritmos de visión de computador, sobre condiciones semi-controladas y sin requerimientos complejos.

A pesar de los avances en el desarrollo de descriptores de video para la cuantificación oculomotora, los métodos existentes son principalmente restringidos a la captura de movimientos oculares sacádicos, descritos como desplazamiento grande de iris, en diferentes tareas de exploración visual. Estos enfoques requieren esfuerzos mayores para describir los comportamientos sutiles en otras variaciones de movimientos oculares. Adicionalmente, se puede considerar que estos

---

<sup>23</sup> SZYMANSKI, Artur KUBIS, Anna y PRZYBYSZEWSKI, Andrezj W.. “Fuzzy Rough sets theory applied to parameters of eye movements can help to predict effects of different treatments in Parkinson’s patients”. En: *International Conference on Pattern Recognition and Machine Intelligence* (2015), págs. 325-334.

<sup>24</sup> KHOSLA, Ajit y KIM, Dongsoo. *Optical Imaging Devices: New Technologies and Applications*. CRC Press, 2015.

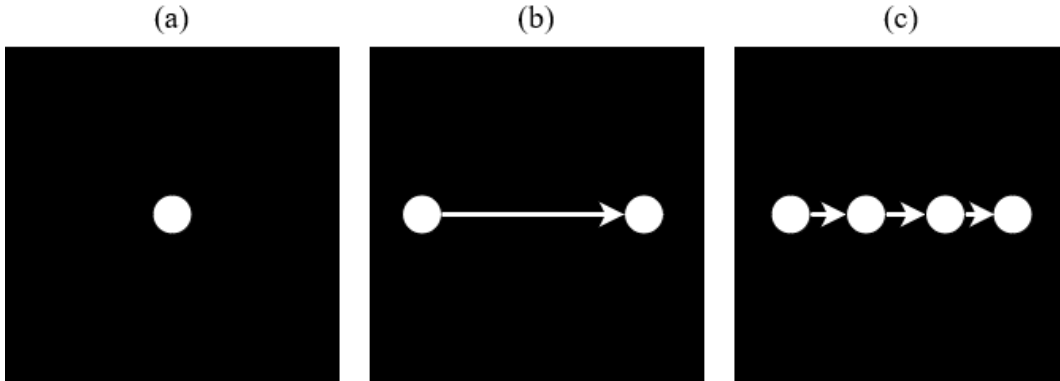
<sup>25</sup> ATEHORTÚAL, Angélica TRUJILLO, David MARTÍNEZ, Fabio. “A characterization of Parkinson’s disease by describing the visual field motion during gait”. En: *11th International Symposium on Medical Information Processing and Analysis* 9681 (2015).

<sup>26</sup> ADHIKARI, Sam y STARK, David E.. “Video-based eye tracking for neuropsychiatric assessment”. En: *Annals of the New York Academy of Sciences* 1387.1 (2017), págs. 145-152.

<sup>27</sup> LAI, Hsin-Yu. “Enabling Saccade Latency Measurements with Consumer-Grade Cameras”. En: *International Conference on Image Processing (ICIP)* 25 (2018), 3169–3173.

movimientos pueden ser cubiertos por movimientos de la cabeza.

**Figura 1.** Ejemplos de movimientos oculares. (a) Fijación ocular. En esta tarea el ojo está fijo en un punto exacto. (b) Movimiento sacádico. En esta tarea el ojo se mueve instantáneamente de un punto al otro. Muy común en tareas como la lectura rápida. (c) Rastreo o búsqueda suave. En esta tarea el ojo hace un seguimiento lento de un punto mientras se mueve de un lugar a otro.



Los patrones de fijación han sido igualmente capturados y caracterizados desde protocolos típicos, siendo aún más desafiante su recuperación debido a el carácter invasivo de algunas técnicas. Además los gestos naturales de la captura pueden insidir en los patrones de temblor que se pretenden caracterizar durante cada experimento. En este sentido, en la literatura se han propuesto métodos alternativos para el modelamiento de patrones de fijación ocular sobre secuencias de vídeo, siendo clasificados automáticamente usando estrategias de aprendizaje de máquina. Una ventaja fundamental en estos esquemas, son los protocolos flexibles de captura desde cámaras convencionales, que además usan descripciones densas del ojo, la cual pueden ser útil para capturar mayor información relacionada con la fijación. Estos enfoques se basan principalmente en modelos discriminatorios, bajo la premisa de clasificación o determinación de la frontera entre dos poblaciones. Estos análisis asumen inicialmente un conjunto significativo de muestras para representar la población, así como también un *prior* con respecto al balance de los datos, lo cual puede ser difícil de asimilar en un contexto clínico.<sup>13</sup>

### 1.3. REDES GENERATIVAS ADVERSARIAS

En este trabajo se intenta encontrar nuevas relaciones espacio-temporales en vídeos de fijación ocular que puedan estar correlacionadas con la EP. Además nosotros intentamos modelar escenarios donde prevalece únicamente la clase de interés, por lo tanto se espera aprender embebidos de forma generativa, suministrado por la clase de interés, es decir, la EP. Es por ello que en esta sección daremos una introducción a las arquitecturas generativas adversarias.

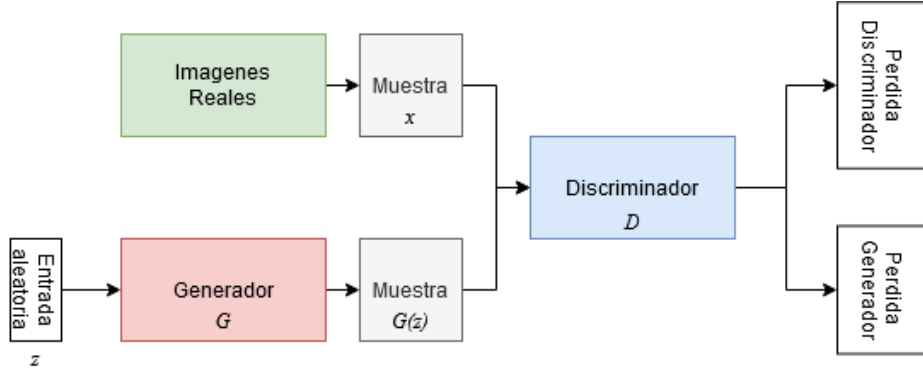
Una red generativa adversaria (GAN, por sus siglas en ingles) es un marco de referencia propuesto por Goodfellow et al. 2014, para estimar modelos generativos por medio de un proceso adversario, en el que se entrenan dos modelos simultáneamente: un modelo generativo  $G$  que captura la distribución de datos, y un modelo discriminatorio  $D$  que estima la probabilidad de que una muestra provino de los datos de entrenamiento en lugar de  $G$ . El proceso de entrenamiento para  $G$  es maximizar la probabilidad de que  $D$  cometa un error. Este marco de referencia corresponde a un juego de mínimos y máximos de  $G$  versus  $D$ . Esta arquitectura ha alcanzado gran éxito en varias tareas de síntesis, tales como traducciones de texto a imagen, traducción de imagen a imagen, el realzamiento de imágenes, entre muchos otros.<sup>28</sup> La arquitectura estándar para las arquitectura de tipo GAN son ilustradas en la figura 2. A continuación se describe cada uno de sus componentes.

- **Generador:** Definido como  $G$ , es generalmente un modelo perceptrón multicapa que captura la distribución de datos. Este modelo aprende a generar datos plausibles a partir de una distribución de espacio latente  $p_z$  de entrada. Su objetivo es minimizar la precisión del discriminador  $D$ , haciendo la información generada más parecida a los datos de la distribución real. Las instancias generadas se convierten en ejemplos de entrenamiento adversarios para el discriminador. Su función es análoga a la de un grupo de falsificadores,

---

<sup>28</sup> MIRZA, Mehdi XU, Bing WARDE-FARLEY, David OZAIR, Serjhil COURVILLE, Aaron GOODFELLOW, Ian POUGET-ABADIE, Jean y BENGIO, Yoshua. “Generative adversarial nets”. En: *Advances in Neural Information Processing Systems* (2014), 2672–2680.

**Figura 2.** Red generativa adversaria. Propuesta por Goodfellow en el 2014, es una red que permite la generación de datos según una distribución real.



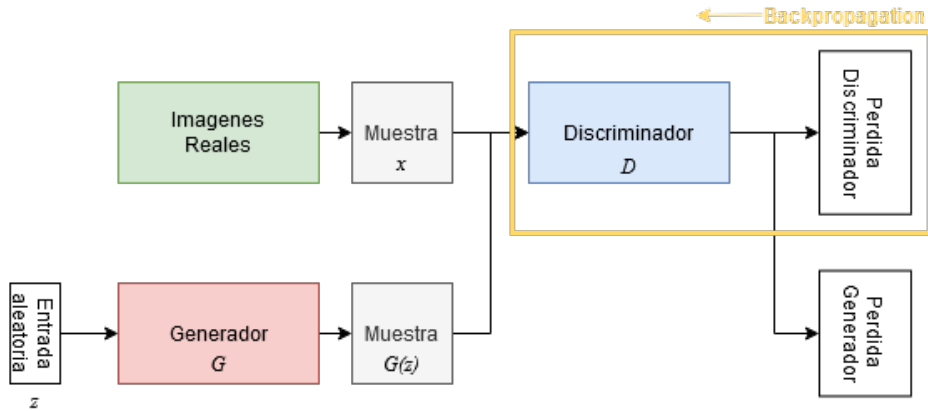
tratando de producir monedas falsas y hacer uso de ellas sin ser detectadas por  $D$ .

- Discriminador:** Definido como  $D$ , es generalmente un modelo perceptrón multicapa que regresa un escalar, representando la probabilidad de que cierto dato pertenezca a la distribución real  $p_z$ . En el mejor caso, se alcanzará un punto óptimo en el entrenamiento de las dos redes, donde el discriminador mapeara cualquier muestra sintética a una probabilidad de  $\frac{1}{2}$ , lo que significa que el discriminador ya no es capaz de distinguir entre la distribución de datos reales  $p_x$  y la distribución de datos falsa  $p_g$ . Su objetivo es maximizar la probabilidad de asignar la etiqueta correcta tanto a ejemplos de entrenamiento como a muestras de  $G$ .

Para entrenar este esquema (ver figura 3), se deben entrenar de manera simultánea ambas redes  $G$  y  $D$ . Para el discriminador  $D$  se sigue el siguiente bucle:

1. El discriminador clasifica los datos reales y falsos del generador.
2. La pérdida del discriminador penaliza al discriminador por clasificar de manera errónea una instancia real como falsa o una instancia falsa como real.
3. El discriminador actualiza sus pesos mediante retro propagación desde la pérdida del discriminador hasta la red del discriminador.

**Figura 3.** Fase de entrenamiento del modelo discriminador.



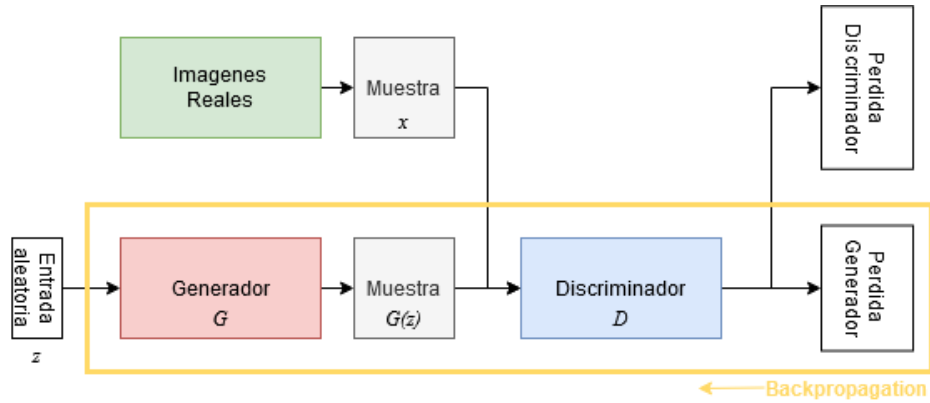
Para el generador  $G$  se sigue el siguiente bucle (ver ilustración 4):

1. Se toma una muestra de espacio latente  $z$  de una distribución  $p_z$
2. El generador produce una salida usando el ruido  $x_g$
3. El discriminador produce una clasificación según la entrada que proviene del generador  $D(x_g)$ .
4. Se calcula la pérdida del discriminador.
5. Se propaga hacia atrás a través del discriminador y el generador para obtener gradientes. Se usan esos gradientes para actualizar los pesos del generador.

Como el discriminador y el generador tienen procesos de entrenamiento diferente, la GAN procede a entrenar por períodos alternantes entre ambas redes.

El generador queda constante durante la fase de entrenamiento del discriminador. A medida de que el entrenamiento del discriminador trata de averiguar como distinguir datos reales de los falsos, tiene que aprender a reconocer las fallas del generador. De igual manera, el discriminador queda constante durante la etapa de entrenamiento del generador. Esta forma de entrenamiento asegura que la red convergerá en un punto.

**Figura 4.** Fase de entrenamiento del modelo generativo.



La función de costo original es derivada de la entropía cruzada entre las distribuciones reales y generadas.

$$\min_G \max_D V(D, G) = \mathbb{E}_{x \sim p_x(x)} [\log(D(x))] + \mathbb{E}_{z \sim p_z(z)} [\log(1 - D(G(z)))]$$

donde

- $D(x)$  es la probabilidad de que la estimación del discriminación de que el dato real de la instancia  $x$  sea real.
- $\mathbb{E}_{x \sim p_x(x)}$  es el valor esperado sobre todas las instancias de los datos reales .
- $G(z)$  es la salida del generador cuando se le es ingresado el ruido  $z$ .
- $D(G(z))$  es la probabilidad de que la estimación del discriminación de que el dato falso sea real.
- $\mathbb{E}_{z \sim p_z(z)}$  es el valor esperado sobre todas las entradas aleatorias al generador.

A partir de la arquitectura estándar GAN han surgido múltiples alternativas que pretenden solucionar problemas específicos o intentan suplir falencias como la discriminación entre instancias de una población o mayor control en la generación de datos. En las siguientes subsecciones

se hace un breve resumen de esquemas generativos adversarios que podrían ser de interés para el presente trabajo.

**1.3.1. GAN CONDICIONAL** En las GANs convencionales propuestas por Goodfellow et al. 2014 , no hay condición sobre la cual entrenar, por lo tanto no hay control sobre las modalidades de datos siendo generados. Sin embargo, si se condiciona el modelo con información adicional es posible dirigir el proceso de generación de datos, discriminando entre instancias en el conjunto de datos que puedan tener algún interés particular.

Las GANs condicionales (CGANs) extienden una GAN estándar a un modelo condicional con alguna información extra  $y$  ya sea en  $G$  o en  $D$ . La condición  $y$  puede ser cualquier información auxiliar, como la etiqueta de clase o datos de otras modalidades. Es posible condicionar el discriminador  $D$  y el generador  $D$  alimentando  $y$  como una capa de entrada adicional. El generador combina el ruido previo  $p_z(z)$  y  $y$  en una representación adjunta oculta, el marco de referencia de entrenamiento adversario permite una flexibilidad considerable en cuanto a cómo se compone esta representación oculta. En el discriminador,  $x$  y  $y$  son representados como entradas a una función discriminadora.<sup>29</sup> La función de costo considerando la condición es:

$$\min_G \max_D V(D, G) = \mathbb{E}_{x \sim p_x(x|y)}[\log(D(x))] + \mathbb{E}_{z \sim p_z(z)}[\log(1 - D(G(z|y)))]$$

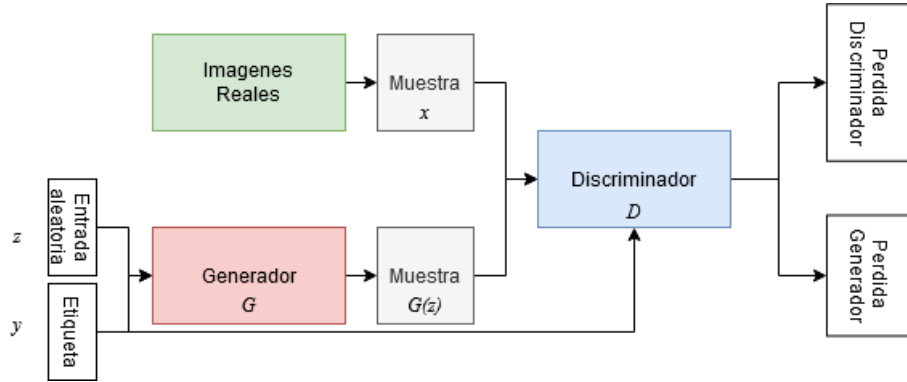
donde  $(\mathbb{E}_{x \sim p_x(x|y)}[\log(D(x))])$  entrena el discriminador para maximizar la probabilidad de los datos de entrenamiento, y  $\mathbb{E}_{z \sim p_z(z)}[\log(1 - D(G(z|y)))]$  entrena tanto al discriminador a minimizar la probabilidad de los datos tomados del generador, y al generador a maximizar la probabilidad de que el discriminador asigne a su propia muestra.

En el discriminador,  $x$  y  $y$  son presentados como entradas a una función discriminadora. En el generador, la entrada de ruido  $z$  y  $y$  son combinados en la representación conjunta oculta. La arquitectura de la CGAN permite la clasificación de los pacientes en el contexto de nuestro

---

<sup>29</sup> MIRZA, M. y OSINDER, S. “Conditional Generative Adversarial Nets”. En: (2014).

**Figura 5.** Arquitectura de la CGAN, que cuenta con etiquetas de clasificación.



problema de fijación ocular, esto es, los pacientes con EP y los sujetos control. Al aprender la representación latente de datos de una clase es posible compararlos con los de la otra clase.

#### 1.4. DETECCIÓN DE ANOMALÍAS

La detección de anomalías es un marco de trabajo estadístico que ha permitido modelar diferentes problemas, en áreas tan diversas como la manufactura, las imágenes médicas y la ciberseguridad. De manera general, los métodos de detección de anomalías se centran en detectar y modelar datos atípicos con respecto a distribuciones normales de una clase específica. Estos datos por su naturaleza suelen reportar una variabilidad muy alta, pueden ser complejos y de alta dimensionalidad.<sup>30</sup>

Esta detección es un problema clásico en áreas como visión por computador, esto es, la determinación de lo normal de lo anormal cuando los datos tienden altamente a ser de una clase (normal) dado el tamaño insuficiente de muestras de la otra clase (anómalo). A pesar de que esto se puede tratar como un problema de aprendizaje supervisado, una aproximación más cercana y realista sería el de detectar el caso anómalo desconocido que conduce al espacio del paradigma

---

<sup>30</sup> LECOAT, B. MANEK, G. ZENATI, H. FOO, C.S. y CHANDRASEKHAR. V.R.. “Efficient GAN-Based Anomaly Detection”. En: (2018).

de aprendizaje de una clase.<sup>31</sup>

Una necesidad común cuando se analizan conjuntos de datos del mundo real es determinar cuáles instancias son atípicas en una población observada. Tales instancias son conocidas como anomalías (o anormalidades, datos desviados, o atípicos); datos que están localizados lejos del volumen principal de datos, y que pueden surgir por varias razones, en diversas aplicaciones, tales como: acciones malintencionadas, fallas de sistema, o fraude intencional. Estas anomalías revelan ideas interesantes sobre los datos y casi siempre traen información valiosa sobre de los mismos.<sup>6</sup>

El objetivo de la detección de anomalías es detectar tales instancias basado en las observaciones sobre un conjunto de datos en particular <sup>32</sup>. Las anomalías pueden ser causadas por errores en los datos, pero a veces son indicativos de un nuevo, y oculto, proceso subyacente. Un dato atípico es una observación que se desvía muy significativamente de otras observaciones de tal manera que levanta sospecha sobre si fue generado por un mecanismo diferente a la distribución típica de los datos <sup>33</sup>. En el campo más amplio del aprendizaje de máquina, los años más recientes han sido testigo de una proliferación de redes neuronales profundas, con resultados sin precedentes a través de varias áreas de aplicación. El aprendizaje profundo es un subconjunto de el aprendizaje de máquina que adquiere buen rendimiento y flexibilidad en aprender a representar información como una jerarquía anidada de conceptos dentro de capas de la red neuronal. En años recientes, los algoritmos basados en aprendizaje profundo para detección de anomalías se han vuelto cada vez más populares y han sido aplicados para diversas tareas.<sup>34</sup>

Las redes generativas adversarias (GANs) y otros procesos de aprendizaje profundo han sido

---

<sup>31</sup> ABARGHOUEI, A.A. ACKAY, S. y BRECKON, T.P.. “GANomaly: Semi-Supervised Anomaly Detection via Adversarial Training”. En: (2018).

<sup>32</sup> BANERJEE, A. CHANDOLA, V. y KUMAR, V.. “Outlier detection: A survey.” En: *ACM Computing Surveys* (2007).

<sup>33</sup> HAWKINS, D. *Identification of Outliers*. Chapman y Hall, 1980.

<sup>34</sup> CHAWLA, S. CHALAPATHY, R.. “Deep learning for anomaly detection: A survey”. En: (2019).

empleados recientemente para resolver este problema, obteniendo resultados destacables.<sup>35 36 32</sup> Por ejemplo, Paffenroth et al. 2018 utilizó análisis robusto de componentes principales (RPCA) para extraer los valores atípicos del dataset de detección de intrusos DARPA en el Lincoln Lab. Zhou et al. 2017 utilizó autoencoders robustos profundos para encontrar anomalías en el dataset MNIST. También Ackay et al. 2018 inventó una arquitectura de GAN novedosa para detección de anomalías en datasets de referencia.

**1.4.1. DETECCIÓN DE ANOMALÍAS CON RPCA** El estudio en el área creciente de Análisis Robusto de Componentes Principales (RPCA) tiene varios algoritmos que se enfocan en la recuperación de subespacios de baja dimensión de datos que han sido corrompidos por información atípica. En particular, hay muchos trabajos que se han tratado en problemas de recuperación donde se demuestran teoremas, y proveen demostraciones matemáticas, en función de asumir que existe una mejor aproximación a un subespacio  $L$  de baja dimensión y una colección verdadera de anomalías  $S$ , se podría recuperar estos valores de su suma  $M = L + S$ , asumiendo que  $L$  y  $S$  cumplen ciertas condiciones.<sup>36</sup>

Pero en contraposición a los problemas de recuperación, el enfoque de *Paffenroth et al. 2018*, se concentra en problemas de detección, en donde teniendo en cuenta el PCA clásico es posible hacer dos preguntas diferentes:

- Dados unos datos (probablemente corruptos por ruido) es posible encontrar un verdadero subespacio de baja dimensión que abarca todos los datos (datos sin alteraciones)?
- Dados unos datos es posible encontrar un subespacio de baja dimensión que sea más efectivo para otro algoritmo de detección, como por ejemplo el PCA es usado como un paso de preprocesamiento para otros algoritmos de aprendizaje de máquina?

---

<sup>35</sup> DE SIMONI, Michelle GHELFI, Emanuel DI MATTIA, Federico GALEONE, Paolo. “A Survey on GANs for Anomaly Detection”. En: (2019).

<sup>36</sup> KAY, K. PAFFENROTH, R. y SERVI, L.. “Robust PCA for Anomaly Detection in CyberNetworks”. En: (2018).

Cuando el PCA es usado como un paso de preprocesamiento para algún algoritmo contra la corriente, es clásico considerar como tanto una dimensión  $k$  objetivo para el subespacio computado, o equivalentemente, un umbral deseado  $\gamma$  para los valores singulares de una matriz de bajo rango  $L$ . Tal umbral es equivalente a la declaración sobre el error máximo posible aceptable entre el dato original y su proyección. De tal manera,  $\gamma$  puede ser escogido basado en un error *a priori*, o en el rendimiento de un *cross-validation* del algoritmo de detección. Los algoritmos de RPCA trabajados en el trabajo de *Paffenroth et al. 2018*, tienen un parámetro  $\lambda$  que controla el intercambio entre la matriz de bajo rango  $L$  y la matriz escasa  $S$ . Imitando al problema de PCA, es posible hacer las siguientes preguntas:

- Dados unos datos (probablemente corruptos por ruido) como debería ser  $\lambda$  para conducir a la mejor recuperación del verdadero subespacio de baja dimensión y la matriz escasa con las anomalías  $S$ , ¿de donde los datos observados fueron construidos?
- ¿Dados unos datos, como es posible hallar una dimensión  $k$  y un valor  $\lambda$  que conduzcan a un subespacio de baja dimensión y un conjunto de anomalías que son efectivas para otro algoritmo de corriente que solo logra entrenar  $\lambda$  y no los numerosos parámetros del algoritmo? En otras palabras, ¿cómo usar RPCA como paso de preprocesamiento para otro algoritmo de aprendizaje de máquina?<sup>28</sup>

**1.4.2. DETECCIÓN DE ANOMALÍAS CON AUTOENCODERS** Un autoencoder profundo es un tipo red neuronal multicapa artificial usada para aprender codificaciones de datos de manera no supervisada en la cual la salida deseada es la misma entrada. De primer vistazo este proceso puede verse trivial ya que la identidad del mapeo puede no tener reconstrucción. Sin embargo, los autoencoders se convierten en no triviales cuando el mapa de identidad es rechazado por ya sea algún tipo de regularización o, más importante, la actual desviación, teniendo capas ocultas las cuales son de una representación no-lineal y de baja dimensión de los datos de entrada. Particularmente, los autoencoders aprenden un mapeo de la entrada hacia sí

mismo a través de un par de fases de codificación y decodificación

$$\hat{x} = D(E(x))$$

donde  $x$  son los datos de entrada,  $E$  es el mapa de codificación de los datos de entrada a la capa oculta,  $D$  es mapa de decodificación de la capa oculta a a la capa de salida y  $\hat{x}$  es la versión recuperada de los datos de entrada. La idea es entrenar  $E$  y  $D$  para minimizar la diferencia entre  $x$  y  $\hat{x}$

Los autoencoders profundos y otras redes neuronales han demostrado su efectividad en descubrir características no lineales, codificadas en vectores embebidos, los cuales resultan descriptores ideales en diversas áreas de aplicación. Sin embargo, en muchos problemas del mundo real, es común encontrar valores atípicos y ruido agudo, y no es muy probable tener acceso a la limpieza de datos de entrenamiento requerido por autoencoders profundos de ruido estándar.<sup>37</sup> En el estado del arte existen trabajos que utilizan los autoencoders profundos que no solo mantienen una calidad alta y características no-lineales pero que también son capaces de eliminar ruido y valores atípicos sin acceder a los datos de entrenamiento. Estos modelos son capaces de dividir la información de entrada en dos partes,  $x = L_D + S$ , donde  $L_D$  puede ser reconstruido por un autoencoder y  $S$  contiene la información atípica en  $x$ . Estos modelos dan credida a enfoques novedosos de detección de anomalías, con rendimiento alto en problemas de referencia.

**1.4.3. DETECCIÓN DE ANOMALÍAS CON GANS** La detección de anomalías usando GANs es un campo de estudio emergente, donde AnoGAN (Anomaly GAN ) fue la primera red en proponer tal concepto. Con el fin de enfrentar los problemas de rendimiento de la AnoGAN surgió un enfoque basado en BiGAN, referido como Detección de anomalías eficiente basado en GAN (EGBAD), el cual superó el tiempo de ejecución de AnoGAN. Recientemente, un enfoque avanzado de una GAN + codificador excedió ambos tiempos de ejecución, siendo

---

<sup>37</sup> ZHOU, C y PAFFENROTH, R. “Anomaly Detection with Robust Deep Autoencoders”. En: (2017).

conocida como Ganomaly.<sup>35</sup> continuación se hace un recuento breve sobre las dos arquitecturas que precedieron las Ganomaly, la cual será descrita en detalle en la metodología propuesta.

- **AnoGAN:** La meta de AnoGAN es el uso de una GAN estándar, entrenada sólo en ejemplos positivos, para aprender un mapeo del espacio latente ( $z$ ) a la instancia real  $\hat{x} = G(z)$  y usar esta representación aprendida para mapear nuevas instancias desconocidas de vuelta al espacio latente.<sup>38</sup> Entrenar una GAN en muestras normales hace que el generador aprenda el colector  $X$  de muestras normales. Dado que el generador aprende como generar muestras normales, cuando una imagen anómala es reconstruida, su reconstrucción será no anómala; de modo que la diferencia entre la entrada y la imagen reconstruida podrá resaltar las anomalías (figura 2.6).
- **EGBAD:** Este modelo trae la arquitectura de la BiGAN a el dominio de la detección de anomalías. En particular, EGBAD trata de resolver las desventajas de AnoGAN que permite aprender un codificador  $E$  que es capaz de mapear muestras de entrada a su representación latente durante el entrenamiento de adversarios.<sup>30</sup>

El puntaje de anomalía se define como:

$$A(x) = \alpha L_G(x) + (1 - \alpha) L_D(x)$$

donde

1.  $L_G(x) = \|x - G(E(x))\|_1$
2.  $L_D(x) = \sigma(D(x, E(x)), 1)$  ( $\sigma$  es la entropía cruzada del discriminador cuando  $x$  es un ejemplo real), or

---

<sup>38</sup> WALDSTEIN, S.M. SCHMIDT-ERFURTH, U. SCHLEGL, T.. SEEBCK, P. y LANGS, G.. “Unsupervised Anomaly Detection with Generative Adversarial Networks to Guide Marker Discovery”. En: (2017).

3. Una pérdida de feature-matching  $L_D(x) = \|f_D(x, E(x)) - f_D(G(E(x)), E(x))\|_1$ , con  $f_D$  regresando la capa precediendo los logits de las entradas dadas en el discriminador.

## 2. PLANTEAMIENTO Y JUSTIFICACIÓN DEL PROBLEMA

La actual adquisición de sistemas para examen oculomotor cuantitativo no representa todo el campo de movimiento ocular y sus deformaciones intrínsecas. Adicionalmente, los métodos de captura, hoy en día, requieren procesos especializados de calibración y montaje, por lo cual su uso es limitado únicamente en escenarios de investigación. Otras alternativas realizan la evaluación oculomotora en video usando cámaras de grado de consumo. Sin embargo, no consideran las alteraciones oculomotoras sutiles que demuestran ser de gran importancia para asistir al diagnóstico temprano de EP y al rastreo de enfermedad progresiva. El problema ha sido ya abordado utilizando vectores embebidos que representan la fijación ocular. Sin embargo, estos patrones establecen condiciones y aprenden distribuciones sobre las clases de entrenamiento. Estas consideraciones pueden limitar el trabajo en escenarios con amplia variabilidad, como en la EP. Además, los trabajos anteriores se limitan a otorgar etiquetas generales sobre las muestras de entradas, lo que reduce el entendimiento de los modelos.

¿Cómo describir anomalías oculomotoras registradas en vídeo que podrían estar asociadas con la enfermedad del Parkinson?

### 3. OBJETIVOS

#### Objetivo general

Implementar una red generativa adversaria para detectar anomalías de micro-movimientos oculares, que permitan soportar el diagnóstico de la enfermedad del Parkinson.

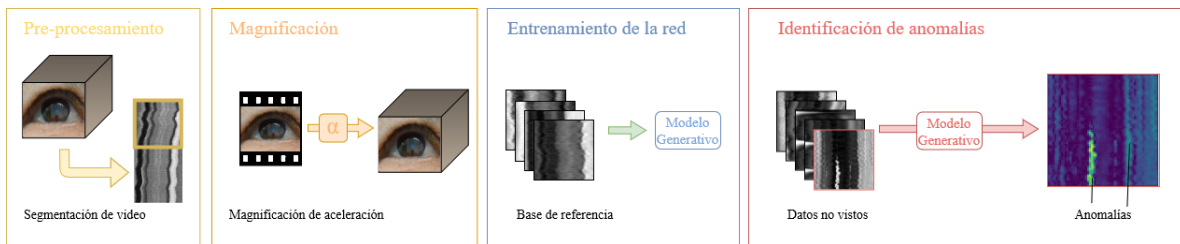
#### Objetivos específicos

- Cuantificar imágenes espacio-temporales que registren patrones de micro-movimiento ocular
- Identificar un esquema de red generativas que permita modelar patrones anormales del Parkinson como un problema de detección anomalías
- Implementar un enfoque generativo para la detección de patrones anormales relacionados con la EP
- Validar el esquema propuesto con respecto a la identificación de anomalías relacionadas con la EP

## 4. REPRESENTACIÓN DE UN BIOMARCADOR DIGITAL PARA LA FIJACIÓN OCULAR USANDO UN ENFOQUE PROFUNDO GENERATIVO

En este trabajo se propuso una metodología de detección de anomalías, implementando y adaptando una red generativa de tipo GANomaly, permitiendo cuantificar y clasificar población diagnosticada con Parkinson, con respecto a la población control. En la Figura 6 se ilustra la metodología general llevada a cabo para el modelamiento del problema planteado. Inicialmente se registran un conjunto de videos de pacientes que observaron un punto fijo. De estos videos son tomados planos transversales al video para capturar micro movimientos relacionados con la fijación ocular. Estos videos son tratados de forma cruda, pero también son amplificados para resaltar diferencias temporales. Luego los datos obtenidos que corresponden a la población de Parkinson son utilizados para aprender un modelo generativo que capture las principales relaciones ocultas que representan estadísticamente esta población. Finalmente, bajo un marco de detección de anomalías se usan ejemplos de las dos secuencias para determinar clasificar las dos poblaciones consideradas. A continuación se describen cada una de estas etapas.

**Figura 6.** Idea general de identificación anómala. Al entrenar una GAN con muestras normales, al alimentarla con información anómala se puede evidenciar una diferencia prudente entre la imagen de entrada y la imagen de salida.



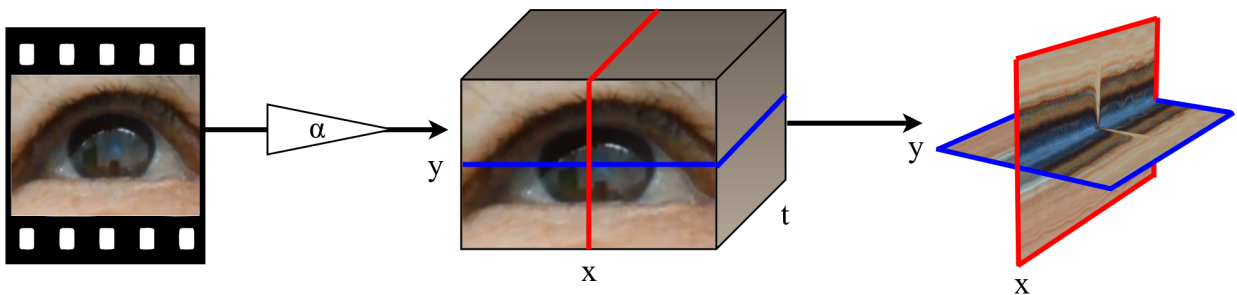
### 4.1. SEGMENTACIÓN DE VIDEOS

Con el fin de proporcionar el modelo segmentos espacio-temporales que codifiquen mapas con los patrones de temblor durante los experimentos de fijación visual, es necesario hallar una

representación de cada video de cada paciente en el dataset. Al obtener un descriptor de video eficiente, el experimento se verá beneficiado. Esta segmentación se logró con el uso de cortes a través del eje del tiempo, por cada video.

Específicamente, se obtuvieron 5 videos de aproximadamente cinco segundo de duración, en el cual se registra la segundos de fijación oculomotora por cada ojo de paciente. Este periodo de tiempo resulta suficiente para registrar patrones rítmicos y periódico del temblor ocular, del cual se conoce que tiene características aproximadas al temblor en otros miembros del cuerpo (alrededor de 6Hz). Por cada video, se hace un corte horizontal y vertical para cada volumen de datos (ver figura 7). Estos cortes logran registrar cambios de intensidad a lo largo del tiempo, para diferentes segmentos del ojo, pudiendo tener una correlación con el temblor ocular. Las posiciones en el eje vertical y horizontal son centradas en el iris, permitiendo capturar de forma simétrica la geometría del ojo, y resaltando los patrones envolventes del temblor ocular. A partir de estos mapas se hace mas fácil la caracterización de patrones oculares, evitando un sesgo de la redundancia de fondo y patrones de fondo, que no tienen relevancia clínica.

**Figura 7.** Segmentación del video. Para cada video se extraen cortes horizontales y verticales.



## 4.2. MAGNIFICACIÓN DE ACELERACIÓN DE VIDEO

En la literatura de visión por computador han emergido técnicas que implementan filtros temporales con el fin de exagerar pequeños movimientos en las secuencias de video. Estos filtros resultan de gran interés porque permiten resaltar comportamientos temporales importantes, por

ejemplo, relacionados con la frecuencias típicas del Parkinson. Es por ello, que en este trabajo se exploró la posibilidad de obtener nuevas representaciones a partir de la magnificación de los videos. Para esto se toman los mismos videos antes de ser segmentados y se les aplica una operación de magnificación de aceleración de video, esto con el fin de ver el movimiento del ojo en presencia del movimiento de la cabeza, para poder analizar los datos de los píxeles locales del ojo.

Un hecho notable sobre las fijaciones de EP es que su comportamiento oscilatorio anormal y sutil el cual ha estado correlacionado fuertemente con la EP <sup>5</sup> <sup>39</sup>. Sin embargo, la cuantificación de estos micro movimientos del ojo constituyen una gran limitación para el análisis del análisis de movimiento. Un desafío adicional esta en la decodificación del desplazamiento menor del ojo cuando está cubierto comparativamente por un movimiento mayor de la cabeza. De modo que en este trabajo empieza por realizar una amplificación espacio-temporal sobre las secuencias de fijación. Un conjunto de frecuencias especificas fueron amplificados usando un enfoque de magnificación de aceleración de video <sup>40</sup>, el cual permite amplificar movimientos sutiles (fijación ocular) cuando están en presencia de movimientos mayores (cabeza). En este trabajo se empleó un esquema de magnificación de aceleración <sup>41</sup>. En esta aproximación se opera en una proyección de la fase de cada cuadro  $I(x, t)$ , en una en una coordenada  $x = [x, y]$ . La fase extraída puede ser representada por una frecuencia espacial  $\omega$  y sus respectivas orientaciones  $\theta$ . Entonces, la La fase local es extraída y descompuesta en una serie de Taylor de segundo orden alrededor de

---

<sup>39</sup> GITCHEL, George. “Experimental support that ocular tremor in Parkinson’s disease does not originate from head movement”. En: *Parkinsonism and related disorders* 20.7 (2014), 743–747.

<sup>40</sup> PINTEA, Silvia ZHANG, Yichao y VAN GEMERT, Jan. “Video Acceleration Magnification”. En: *Computer Vision and Pattern Recognition* (2017).

<sup>41</sup> FLEET, David J.and JEPSON, Allan D.. “Computation of component image velocity from local phase information”. En: *International Journal of computer vision* 5.1 (1990), 77–104.

$t$  <sup>42</sup>:

$$\phi_{\omega,\theta}(x, t + 1) = \phi_{\omega,\theta}(x, t) + \frac{\partial\phi_{\omega,\theta}(x, t)}{\partial t} + \frac{1}{2} \frac{\partial^2\phi_{\omega,\theta}(x, t)}{\partial t^2} \quad (1)$$

El termino de primer orden en la ecuación 2 representa cambios de magnitud lineal en la velocidad, mientras que el término de segundo orden mide la desviación del cambio de movimiento, es decir, la aceleración. La velocidad cuantifica todos los cambios de movimiento sin discriminar entre los grandes y los pequeños. En cambio en el componente de la aceleración se pueden amplificar los movimientos menores en presencia de las movimientos grandes. Este hecho asume, desde el punto de vista de la aceleración, que todas los movimientos mayores son aproximadamente lineales en la escala temporal de los movimientos menores. Toda los componentes lineales son ignoradas en el análisis y por lo tanto solo se toma en cuenta la derivada de segundo orden, la cual es obtenida a través del Laplaciano del operador Gaussiano:

$$\phi_{\omega,\theta}^n(x, t) = \frac{\partial^2\phi_{\omega,\theta}(x, t)}{\partial t^2} = \phi_{\omega,\theta}(x, t) * \frac{\partial^2 G_\sigma(x, t)}{\partial t^2} \quad (2)$$

donde  $\sigma$  es el filtro Gaussiano estándar de desviación. Entonces, las fases son ampliadas como:

$$\phi_{\omega,\theta}(x, t) = \phi_{\omega,\theta}(x, t) + \alpha\phi_{\omega,\theta}^n(x, t) \quad (3)$$

donde  $\alpha$  es el factor de magnificación, el cual amplifica los cambios de segundo orden a una frecuencia de  $f = \frac{f_r}{8\pi\sqrt{2}\sigma}$ , donde  $f_r$  es el frame rate del video.

Al realizar esto, se realiza la segmentación del video de este nuevo dataset usando el procedimiento de la anterior sección.

---

<sup>42</sup> PORTILLA, Javier y SIMONCELLI, Eero. “A parametric texture model based on joint statistics of complex wavelet coefficients”. En: *Internacional Journal of computer vision* 40.1 (2000), 49–70.

### 4.3. ARQUITECTURA GANOMALY PARA REPRESENTAR PATRONES PARKINSONIANOS

Un objetivo fundamental en este trabajo fue la cuantificación de los patrones del Parkinson presentes en los movimientos de fijación ocular. Por lo tanto el enfoque presentado en este trabajo representa una alternativa al diseño de biomarcadores digitales de la EP utilizando el aprendizaje de varios inter-ejemplos de la EP en contraste a otra variación diferente de la parte aislada del Parkinson. Similarmente, las variaciones de EP sirven como nuestros datos anormales. Siguiendo esto, utilizamos un esquema de detección de anomalías dedicado a aprender solo los patrones de Parkinson, donde cualquier otro comportamiento de fijación oculomotora es identificada como una anomalía.

Los datos de la EP sirvieron como nuestra base de referencia. Entrenamos el modelo generativo haciendo que aprendiera las variaciones presentes en el movimiento ocular de fijación en pacientes con EP. Luego utilizamos este modelo entrenado para representar la parte aislada de los datos obtenidos de los pacientes control. Al hacer este proceso, somos capaces de identificar las principales diferencias entre ambas distribuciones de datos. Esta metodología será nuestro principal biomarcador, que nos permite diferenciar entre las dos poblaciones e identificar patrones de la enfermedad.

En este caso, el objetivo es modelar un conjunto de muestras utilizando datos de una sola distribución, y al mapear cualquier muestra cuyo resultado este por fuera de los límites establecidos por esta distribución se tome como una anomalía. Encontrar y cuantificar estas distancias es algo bastante complejo, y para manejar esta complejidad nos basamos en una arquitectura base de una red GAN estado del arte denominada Ganomaly. LA configuración de esta arquitectura es re-entrenada para obtener una nueva representación dedicada a describir patrones relacionados con la fijación ocular.

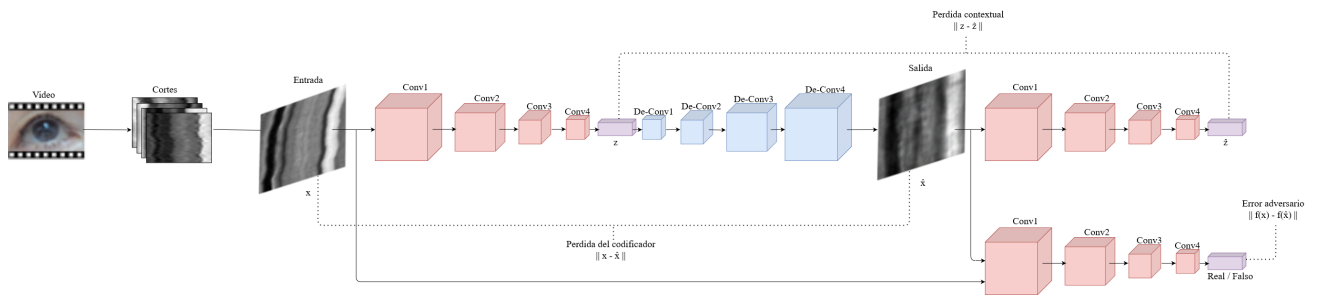
La arquitectura Ganomaly (ver figura 8) es una red generativa adversaria sin supervisar cuyo objetivo es detectar anomalías en un dataset que es altamente sesgada hacia una clase particular, por ejemplo comprendiendo ocurrencias normales no-anómalas para entrenamiento solamente.

La definición formal es la siguiente<sup>31</sup>:

En este contexto se considera un conjunto de datos  $D$  que comprende solo de imágenes normales (de la clase con mayor cantidad de datos en el escenario simulado)  $M$ ,  $D = \{X_1, \dots, X_M\}$  y un dataset mucho más pequeño para entrenamiento  $\hat{D}$  de imágenes  $N$  normales y anormales,  $\hat{D} = \{(\hat{x}_1, y_1), \dots, (\hat{x}_N, y_N)\}$ , donde  $y_i \in [0, 1]$  denota la etiqueta de la imagen. En la configuración práctica, el conjunto de entrenamiento es mucho más grande que el conjunto de prueba.

Dado el dataset, el primer objetivo es modelar a  $D$  para aprender sus características, luego detectar las muestras anormales  $\hat{D}$  como parte aislada durante la fase de interferencia. El modelo  $f$  aprende la distribución de datos y minimiza el puntaje de anomalía  $A(x)$ . Para una imagen de prueba  $\hat{x}$ , un puntaje de anomalía alto  $A(\hat{x})$  indica posibles anomalías dentro de la imagen. El criterio de evaluación para esto es un umbral ( $\theta$ ), donde  $A(\hat{x}) \leq \theta$  indica una anomalía.

**Figura 8.** Arquitectura de la red propuesta. Basada en Ganomaly, es una red generativa que es eficaz para tratar datos sesgados.



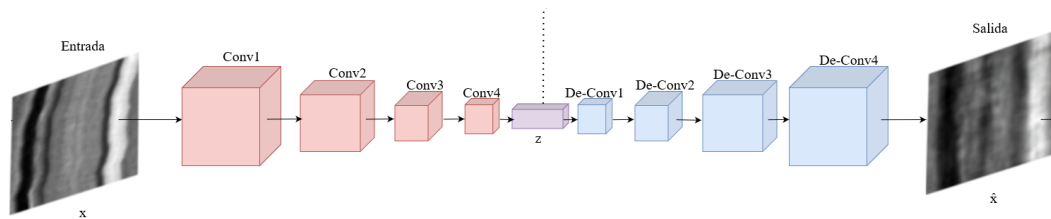
La red Ganomaly consiste de dos redes codificadoras, una red decodificadora y una red discriminadora, empeladas a través de tres subredes. A continuación se describe cada uno de los componentes de la red implementada para el modelamiento del Parkinson desde un marco de trabajo de detección de anomalías.

**4.3.1. SUBRED GENERADORA** La primera subred considerada en este esquema es un autoencoder que se comporta como la parte de generador del modelo. El generador aprende la representación de datos de entrada y reconstruye la imagen entrada a través de una red de codificador y decodificador, respectivamente. La parte decodificadora  $G_D$  de la red del generador

$G$  adopta la arquitectura de un generador DCGAN. Este enfoque incrementa la dimensión del vector  $z$  para reconstruir la imagen  $x$  como  $\hat{x}$ . En otras palabras,  $G$  genera una imagen  $\hat{x}$  a través de  $\hat{x} = G_D(z)$ , donde  $z = G_E(x)$ .

Esta es la primera red por donde empieza la generación de la arquitectura (figura 9). Como esta compuesta por un codificador y un decodificador, la red obtendrá tanto una representación latente tanto como una imagen generada que luego sera entregada la siguiente subred decodificadora. De las dos imágenes involucradas se obtiene un puntaje de diferencia conocido como la perdida del codificador.

**Figura 9.** Subred generadora. Esta red toma la imagen entrante y la codifica en un vector embebido. Luego, decodifica este vector en una nueva imagen reconstruida.

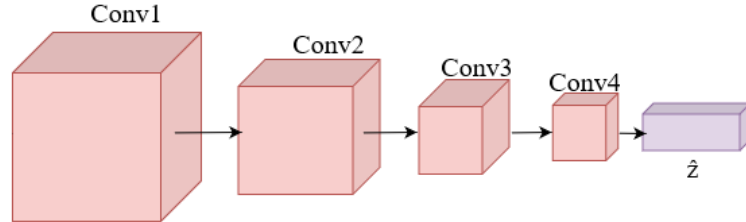


**4.3.2. SUBRED CODIFICADORA** La segunda subred (ver figura 10) es la red del codificador  $E$  que comprime la imagen  $\hat{x}$  que es reconstruida por la red  $G$ .  $E$  reduce la dimensión de  $\hat{x}$  para encontrar su representación de características  $\hat{z} = E(\hat{x})$ .

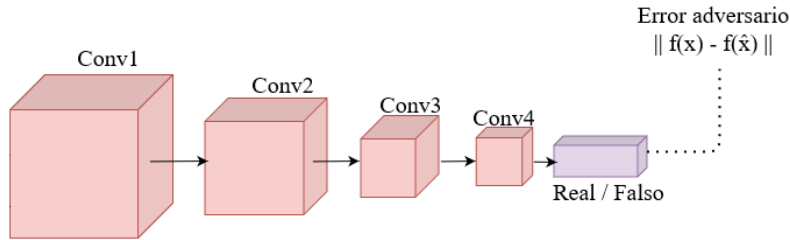
La imagen generada en la subred anterior es el dato de entrada de esta subred. Al obtener la representación latente  $\hat{z}$ , la subred obtiene un puntaje de perdida generado por la diferencia de la representación latente  $z$  y  $\hat{z}$

**4.3.3. SUBRED DISCRIMINADORA** La tercera subred (ver figura 11) es la red del discriminador  $D$  cuyo objetivo es clasificar la entrada  $x$  y la salida  $\hat{x}$  como real o falsa, de acuerdo a la calidad en la reconstrucción obtenida. A esta red entran tanto la imagen original de la distribución como la imagen generada de la subred generadora. De estas dos imágenes se obtiene un puntaje de clasificación como en las redes generativas adversarias estándares .

**Figura 10.** Subred codificadora. Esta red toma la imagen generada y la transforma en un vector embebido.



**Figura 11.** Subred discriminadora. Esta tercera subred es la que da la probabilidad de que la imagen provenga de la distribución real.



#### 4.4. ENTRENAMIENTO DEL MODELO GENERATIVO GANOMALY

Para entrenar el modelo, hacemos la hipótesis de que cuando una imagen anormal es pasada a la red  $G$ ,  $G_D$  no es capaz de reconstruir la anormalidad a pesar de que  $G_E$  logra mapear la entrada  $X$  a espacio latente  $z$ . Esto es porque la red solo está modelada con muestras normales durante el entrenamiento, y su parametrización no es apta para generar muestras anormales. Una salida  $\hat{X}$  que ha perdido sus anomalías puede conducir a la red codificadora  $E$  a mapear  $\hat{X}$  a un vector  $\hat{z}$  que también ha perdido su representación de características anormales, causando una diferencia entre  $z$  y  $\hat{z}$ . Cuando exista esa disparidad dentro del espacio latente para una imagen de entrada  $X$ , el modelo clasifica a  $X$  como una imagen anómala.

Para validar esta hipótesis, se formula nuestra función objetivo combinando tres funciones de pérdida, las cuales optimizan sus respectivas subredes.

- **pérdida adversaria:**  $\mathcal{L}_{adv} = \mathbb{E}_{x \sim p_x} \|f(x) - \mathbb{E}_{x \sim p_x} f(G(x))\|_2$

- **pérdida contextual:**  $\mathcal{L}_{con} = \mathbb{E}_{x \sim px} \|x - G(x)\|_1$
- **pérdida del codificador:**  $\mathcal{L}_{enc} = \mathbb{E}_{x \sim px} \|G_E(x) - E(G(x))\|_2$

El proceso para realizar las pruebas es el siguiente: para una muestra de prueba  $\hat{x}$ , el puntaje de anomalía  $A(\hat{x})$  o  $s_{\hat{x}}$  es

$$A(\hat{x}) = \|G_E(\hat{x}) - E(G(\hat{x}))\|$$

Para el rendimiento total, se computa el puntaje de anomalía de una muestra individual  $\hat{x}$  dentro del conjunto de pruebas  $\hat{D}$ , o cual otorga un puntaje de anomalía  $S = \{s_i : A(\hat{x}), \hat{x}_i \in \hat{D}\}$ . Luego se aplica escalamiento de características para tener el puntaje de anomalía entre el rango de probabilidades  $[0,1]$ .

$$s'_i = \frac{s_i - \text{mín}(S)}{\text{máx}(S) - \text{mín}(S)}$$

Finalmente, la última ecuación regresa un vector de puntaje de anomalía para la evaluación final de todo el conjunto de pruebas.

## **5. DISEÑO EXPERIMENTAL**

### **5.1. CONJUNTO DE DATOS**

Para este estudio se obtuvieron videos de patrones de moción de 13 pacientes Parkinson en total (edad promedio de 72.3) y 13 pacientes control (edad promedio de 72.2).

Los participantes fueron invitados a observar un foco proyectado en una pantalla con un fondo oscuro. Una cámara convencional, Nikon D3200, con una resolución espacial de 1280x720 píxeles y una resolución temporal de 60 FPS fue fijada en frente de los sujetos para capturar la cara entera. La región del ojo fue manualmente cortada (210x140 píxeles) para obtener las secuencias de interés. Un total de 130 secuencias fueron grabadas por cada ojo, es decir, 5 muestras por persona, con una duración de 5 segundos cada una.

Cada video representa un volumen de datos (ver figura 7). Para representar cada video, se extrajeron un corte vertical y horizontal a través del centro de cada volumen, resultando en 260 imágenes de sujetos control y 260 imágenes de pacientes Parkinson.

Este estudio fue aprobado por el Comité de Ética de la Universidad Industrial de Santander en Bucaramanga, Colombia. Los participantes fueron reclutados de la fundación local para el Parkinson FAMPAS (Fundación del Adulto Mayor y Parkinson Santander) y la institución anciana Asilo San Rafael. El consentimiento escrito fue obtenido para cada participante.

### **5.2. CONFIGURACIÓN DE LA ESTRATEGIA**

Antes de empezar con la clasificación de anomalías, se entrena la red con nuestra base de referencia, los cortes de video de los pacientes EP. Al aumentar los datos utilizando traslación, rotación, re-escalamiento y ruido, se obtuvieron 800 imágenes de entrenamiento. Para configurar la red Ganomaly, se usaron 60 iteraciones de epoch, un batchsize de 1, y pesos iniciales de 1, 50 y 1 en las subredes codificadoras, contextuales y adversarias respectivamente. Luego de que la fase de entrenamiento esta terminada, se empezaran las imágenes de la parte aislada, en este

caso nuestros datos de los pacientes control, para encontrar las diferencias en las variaciones. Para la fase de validación, se usaron 100 imágenes de pacientes Parkinson y 207 imágenes de pacientes control. Acá, se obtiene un puntaje de anomalía para cada subred y imagen de prueba, la cual se utilizará para cuantificar el resultado del experimento. Para esta validación, se realizaron experimentos usando un k-fold de 4, alternando los 10 pacientes que fueron usados para la fase de entrenamiento y los 3 pacientes usados para el proceso de validación. Este experimento también se realizó cambiando las bases de referencia en la red. Es decir, cambiando los datos de entrenamiento a los de paciente control, y la parte aislada a pacientes Parkinson. Este experimento se denominó como Ganomaly invertido. También se realizó un experimento extra con la red previamente mencionada AnoGAN para proveer un punto de referencia para la red Ganomaly. La red AnoGAN consta de una estructura menos robusta (no contiene subredes) de una red generativa adversaria estándar para detectar anomalías.

### 5.3. MÉTRICAS DE VALIDACIÓN

A continuación se hace una descripción de las métricas de validación que fueron tenidas en cuenta en el presente trabajo para cuantificar los alcances del método propuesto. Para estas métricas de análisis estadístico de clasificación binaria, se tienen en cuenta cuatro términos, que indican una contabilización de las predicciones para clasificación binaria:

- **Verdaderos Positivos (VP):** Son los resultados que se predijeron correctamente como positivos. En el caso de este experimento corresponde a todos los cortes de video de la base de referencia que fueron clasificados como de la misma base.
- **Falsos Positivos (FP):** Son los resultados que se predijeron incorrectamente como negativos. En el caso de este experimento corresponde a todos los cortes de video de la base de referencia que fueron clasificados como parte aislada.
- **Verdaderos Negativos (VN):** Son los resultados que se predijeron correctamente como negativos. En el caso de este experimento corresponde a todos los cortes de video de la

parte aislada que se clasificaron como parte aislada.

- **Falsos Negativos (FN):** Son los resultados que se predijeron incorrectamente como negativos. En el caso de este experimento corresponde a todos los cortes de video de la parte aislada que fueron clasificados como base de referencia.

De estos términos es posible calcular las siguientes métricas en el análisis estadístico:

- **Precisión:** La precisión, o valor predicho positivo, es la fracción de instancias relevantes entre las instancias recuperadas. Esta medida es la proporción de identificaciones positivas que fueron verdaderamente correctas. Esta dada por

$$\text{Precision} = \frac{VP}{VP + FP}$$

- **Recall:** El recall (o sensibilidad) es la fracción de instancias relevantes que fueron recuperadas. Es la proporción de instancias que son verdaderamente positivas que fueron identificadas correctamente. Esta dado por

$$\text{Recall} = \frac{VP}{VP + FN}$$

- **F1-score:** En el análisis estadístico de clasificaciones binarias, el F1-Score (Puntaje F1) es la medida de precisión de una prueba. Es calculado de el Recall y Precisión de una prueba. Este puntaje está dado por

$$F_1 = 2 \cdot \frac{\text{precision} \cdot \text{recall}}{\text{precision} + \text{recall}} = \frac{VP}{VP + \frac{1}{2}(FP + FN)}$$

- **Curva ROC (Returning Operating Characteristic):** Una curva ROC es un gráfico que muestra el rendimiento de un modelo de clasificación en todos los umbrales de clasificación. Esta curva tiene dos parámetros:

- **Taza de verdaderos positivos (TVP):** Es el mismo Recall, el cual es está dado por

$$\text{TVP} = \frac{VP}{VP + FN}$$

- **Taza de falsos positivos (TFP):** Esta medida esta definida como

$$\text{TFP} = \frac{FP}{FP + VN}$$

La curva ROC traza la relación TVP vs. TFP en diferentes umbrales de clasificación. Al disminuir este umbral se obtendrían un mayor número de verdaderos positivos y falsos positivos, de igual modo que al aumentar el umbral disminuiría esta medidas. Para hallar el ROC de una clasificación, se utiliza la medida de área bajo la curva (AUC). Esta medida AUC provee una medida agregada del rendimiento a través de todos los umbrales de clasificación. Este resultado varía entre (0-1).

## 6. EVALUACIÓN Y RESULTADOS

A continuación se ilustran los resultados obtenidos por la estrategia de detección de anomalías, usando como base una arquitectura de tipo Ganomaly. El experimento de mayor interés en nuestro trabajo, es el uso de ejemplos de Parkinson como datos control, para aprender una representación mono-clase con estos datos y luego discriminar como muestras anómalas los datos correspondientes a sujetos no-Parkinson. En la tabla 1 se resumen los resultados cuantitativos obtenidos para esta representación de Ganomaly, entrenando sobre muestras de pacientes diagnosticados con la enfermedad. Como línea base, el enfoque propuesto se comparó con otra arquitectura del estado del arte, denominada AnoGAN, así como también utilizando la representación Ganomaly, pero esta vez utilizando los ejemplos control para el entrenamiento (Ganomaly invertida). Para la red Ganomaly, se obtuvo un F1 Score promedio de 0.76, el cual es una medida de nuestra precisión de prueba, con un máximo de 1.0. Esta medida fue calculada con una Precisión del 0.625, el cual es el número de resultados positivos verdaderos dividido por el número de todos los resultados positivos, incluyendo los no identificados correctamente, y un Recall del 0.965, el cual es el número de resultados positivos verdaderos dividido por el número de muestras que debieron haber sido identificados como positivos.

**Tabla 1.** Resultados de los tres experimentos representados en promedios. Para cada experimento se realizó un k-fold de 4. El mejor experimento fue el de Ganomaly con la base de referencia EP.

|                    | ROC         | Recall      | Precision   | Threshold   | F1-Score    |
|--------------------|-------------|-------------|-------------|-------------|-------------|
| AnoGAN             | 0.52        | 0.48        | 0.35        | 0.05        | 0.56        |
| Ganomaly invertida | 0.86        | 0.89        | 0.65        | 0.125       | 0.75        |
| <b>Ganomaly</b>    | <b>0.95</b> | <b>0.96</b> | <b>0.62</b> | <b>0.06</b> | <b>0.76</b> |

En un análisis más detallado se graficó la curva ROC para determinar el comportamiento del enfoque propuesto. Esto otorgó una área bajo la curva promedio ROC (Receiver Operating Characteristic) de  $0,95 \pm 0,03$  (ver figura 12). Esto significa que para cada fold del experimento se obtuvo una clasificación eficiente durante el periodo de validación. Del análisis de la curva

**Tabla 2.** Resultados del análisis estadístico usando ANOVA. Ambos enfoques de Ganomaly son capaces de negar la hipótesis nula ya que tienen un valor  $p \leq 0,05$ , mientras que AnoGAN no. Como AnoGAN no tiene subredes, el análisis solo se hace una vez.

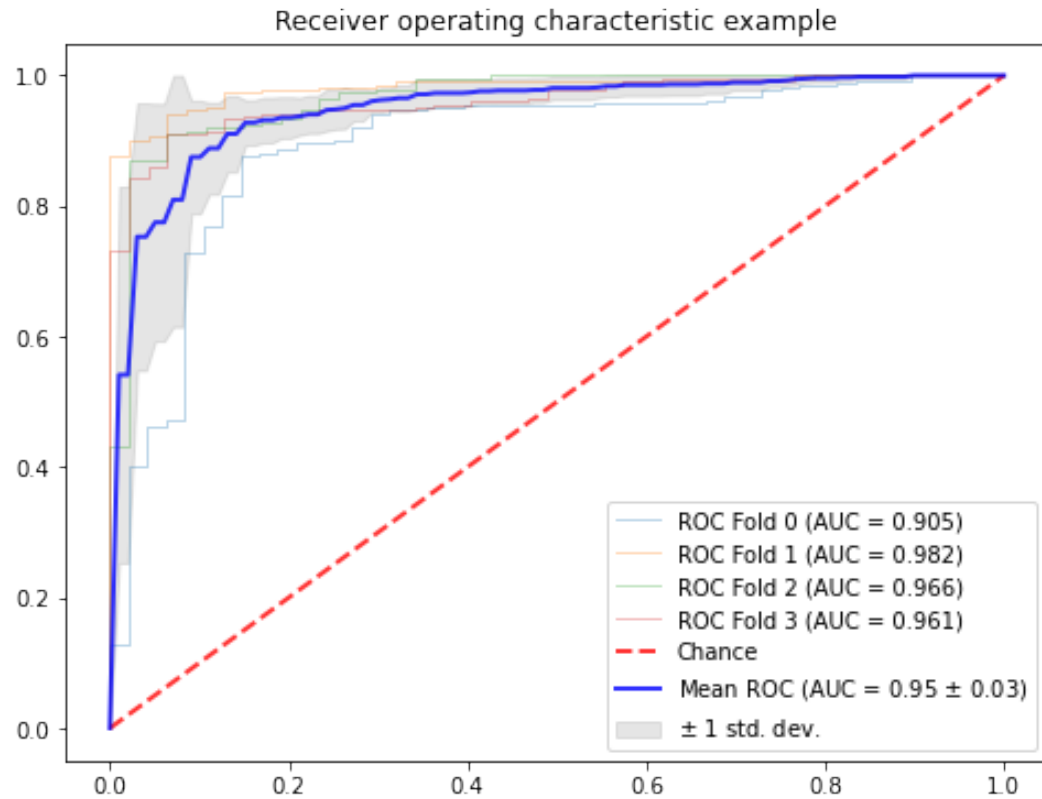
|                   | AnoGAN       | Ganomaly         | Ganomaly (inverted) |
|-------------------|--------------|------------------|---------------------|
| Encoder Error     | p-value=0.65 | p-value=1.34e-16 | p-value=4.27e-11    |
| Contextual Error  | -            | p-value=6.81e-41 | p-value=2.61e-17    |
| Adversarial Error | -            | p-value=0.0038   | p-value=0.023       |

ROC y la comparación de la relación de precisión-recall de ella, se obtuvo un umbral de puntaje anómalo del 0.056, el cual optimizo la clasificación de los pacientes. Esto quiere decir que los pacientes por debajo del umbral pertenecen a nuestra base de referencia (EP), y el resto pertenecen a nuestra parte aislada (control).

Los resultados obtenidos también fueron resumidos en una matriz de confusión, como se ilustra en en la figura 14. De los 3 pacientes EP que se usaron en el proceso de validación, se predijeron 2.64 correctamente. De los 13 pacientes control validados, 11.8 de ellos fueron predichos correctamente, mientras que 1.2 fueron predichos incorrectamente. Este resultado es significativo, teniendo en cuenta la capacidad de representación de patrones Parkinsonianos, logrando una tasa de efectividad del 88 %, así como también una capacidad para discriminar anomalías (ejemplos control que no ha visto la red) de hasta un 91 %. Este resultado evidencia la capacidad del método propuesto como biomarcador digital para discriminar entre dos conjuntos poblacionales, y sobre todo para determinar descriptores de la enfermedad, construidos en los espacios latentes de la representación profunda.

Teniendo en cuenta nuestro objetivo de modelar y caracterizar la enfermedad, en la Figura 13 se ilustra un boxplot para observar las diferencias de ambas distribuciones en cada una de las subredes. En este caso cada imagen es representada por el error de reconstrucción medido en los embebidos de la arquitectura, es decir el error en el codificador, el error adversario y el error contextual. De esto, se realizó un análisis estadístico más detallado. Para el experimento, se

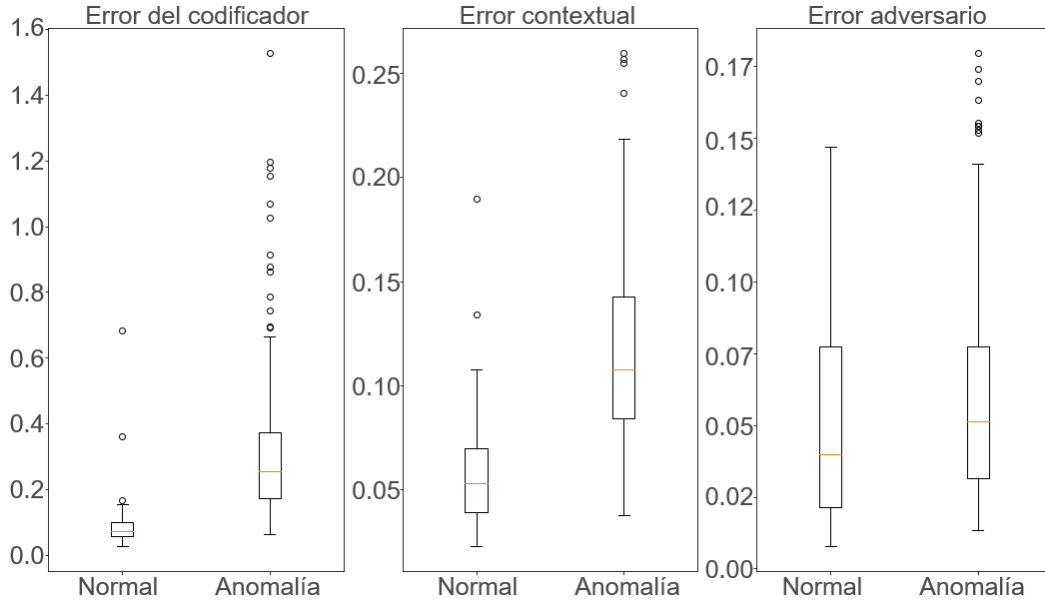
**Figura 12.** Curva ROC del experimento Ganomaly estándar. Como se obtuvo un ROC promedio del 95 %, el modelo es capaz de lograr una validación robusta, nunca sacrificando ni precisión y sensibilidad.



obtuvieron los valores de media y varianza de las poblaciones normales y anormales para cada subred, y luego se realizó un ANOVA de una vía (ver Tabla 2) , con la hipótesis nula siendo aquella de similitud entre ambas poblaciones. Si el valor p en la prueba es menor a 0.05, entonces esta hipótesis será rechazada por la subred analizada, lo cual significa que el enfoque es capaz de notar diferencias significantes entre ambas poblaciones. De esto, se sigue que el enfoque de Ganomaly rechaza la hipótesis nula, mientras que AnoGAN no. De tal manera que el enfoque Ganomaly es capaz de diferenciar entre las poblaciones control y Parkinson.

Adicionalmente, se obtuvieron las subtracciones de algunos ejemplos de imágenes. Al realizar la resta entre las imágenes reales (antes de entrar a la red) y las imágenes falsas (después de entrar a la red) es posible evidenciar una diferencia entre lo que aprendió a representar la red y lo que no obtuvo en su representación latente observando la imagen en su dominio de frecuencias

**Figura 13.** Boxplot del experimento. Para cada subred de la Ganomaly se pueden observar varianzas significativas.

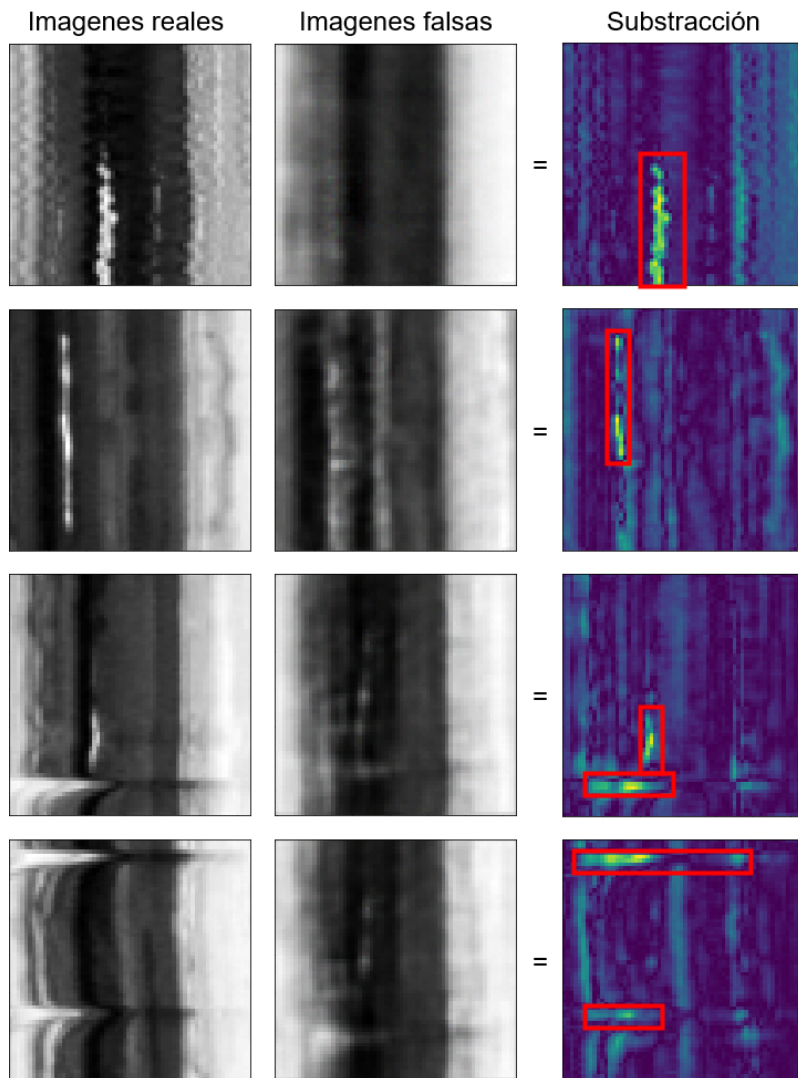


**Figura 14.** Matriz de confusión del experimento. La primera fila representa los pacientes con EP, y la segunda fila representa los pacientes control. Los pacientes en la diagonal son aquellos que fueron clasificados correctamente, mientras el resto son clasificados incorrectamente. Cada bloque contiene información acerca de los cortes de imagen y los pacientes.

|           |                                                                           |                                                                             |
|-----------|---------------------------------------------------------------------------|-----------------------------------------------------------------------------|
| Parkinson | <b>Cortes</b><br>$42.25 \pm 2.58$<br>(88%)<br><br><b>Pacientes</b><br>2.6 | <b>Cortes</b><br>$5.75 \pm 2.58$<br>(12%)<br><br><b>Pacientes</b><br>0.4    |
|           | <b>Cortes</b><br>$24.25 \pm 2.58$<br>(9%)<br><br><b>Pacientes</b><br>1.2  | <b>Cortes</b><br>$235.75 \pm 2.58$<br>(91%)<br><br><b>Pacientes</b><br>11.8 |
|           | Predicho como Parkinson                                                   | Predicho como Control                                                       |

(ver figura 15).

**Figura 15.** Substracción de algunas imágenes antes y después de entrar a la red. Las diferencias son mostradas en el dominio de frecuencias, con las alteraciones en frecuencia relevantes representadas en alto contraste.



## 7. CONCLUSIONES Y PERSPECTIVAS

En este trabajo se introdujo un esquema profundo generativo, diseñado en un marco de detección de anomalías, para la representación y caracterización de patrones parkinsonianos en experimentos de fijación ocular. El método propuesto demostró tener las capacidades de biomarcador digital de la enfermedad, logrando modelar segmentos de video, que visualizan el eje temporal de los micro-movimientos oculares. La representación propuesta se fundamenta en la arquitectura Ganomaly, que demostró un relevante resultado en la detección anómala. Sin embargo, se encontró que en el proceso de entrenamiento se favorecen los experimentos con altos volúmenes de datos de entrada. Por esta razón, se encontró la obligación de usar métodos de aumento de datos para expandir el volumen de los datos que entran a las capas de entrada con el fin de incrementar la precisión de la clasificación.

Durante los esquemas de validación, se implementó una validación de clasificación automática, teniendo en cuenta la detección de anomalías y diferencias descritas desde los embebidos de reconstrucción de la arquitectura implementada y entrenada. La arquitectura implementó un resultado de prueba del 0.935, mucho más allá que el de la implementación de AnoGAN. Con un claro rechazo de la hipótesis nula, el esquema de Ganomaly es capaz de aprender y diferenciar entre dos poblaciones, sin tener que ajustar una distribución paramétrica sobre los datos. Esto sirvió como una identificación de los segmentos espacio-temporales que están asociados con la EP.

Durante la experimentación, se encontró que al usar la población de los pacientes afectados con EP como nuestra base de referencia en el proceso de aprendizaje fue mucho más efectivo en la validación que la población control. Al usar el enfoque de Ganomaly invertido, se logró una diferencia más grande, pero con resultados estadísticos más bajos. Este resultado evidencia la capacidad de la arquitectura para modelar patrones parkinsonianos, los cuales presentan una variabilidad admisible, con respecto al resto de posible sujetos etiquetados como control, los cuales no tienen una afectación específica. En términos de implementación en un escenario

clínico, este hecho resulta relevante, teniendo en cuenta que en centros especializados de atención de pacientes con Parkinson se tendrán muestras diarias de estos pacientes, mientras que los sujetos control no podrán fácilmente ser considerados. Además, una vez entendida la capacidad de discriminación se pueden estudiar representaciones latentes e intermedias que puedan explicar nuevas relaciones complejas asociadas con la enfermedad.

Este experimento de hallazgo de biomarcadores sirvió como otro argumento a favor del temblor ocular vs. temblor de cabeza propuesto por Glitchel, ya que hemos sido capaces de encontrar biomarcadores usando información exclusivamente de fijación oculomotora. Como demostramos que esta arquitectura se beneficia de la cantidad de datos proveídos en el proceso de entrenamiento, se sigue que este podría alcanzar mayores resultados y precisión más alta si se obtuvieran más datos.

El trabajo presente tiene un gran potencial para implementarse como herramienta de soporte en el diagnóstico temprano de la EP usando equipo para el análisis oculomotor. Trabajos futuros incluyen la detección y validación de anomalías en otras tareas de movimiento ocular. Entre ellas están el movimiento sacádico ocular, la búsqueda suave, y el rastreo ocular. También están el análisis individual para cada etapa de la enfermedad del Parkinson, para de esta manera dar mayor evidencia de validación en etapas tanto tempranas como etapas avanzadas. En la etapa de pruebas también incluir nuevas arquitecturas de redes generativas profundas.

## BIBLIOGRAFÍA

ABDULHAY, Enas. “Gait and tremor investigation using machine learning techniques for the diagnosis of Parkinson disease”. En: *Future Generation Computer Systems* 83 (2018), págs. 366-373 (vid. pág. 11).

ACKAY, S. , ABARGHOUEI, A.A. y BRECKON, T.P.. “GANomaly: Semi-Supervised Anomaly Detection via Adversarial Training”. En: (2018) (vid. págs. 24, 37).

ADHIKARI, Sam y STARK, David E.. “Video-based eye tracking for neuropsychiatric assessment”. En: *Annals of the New York Academy of Sciences* 1387.1 (2017), págs. 145-152 (vid. pág. 16).

AJAY, Jerry. “A pervasive and sensor-free deep learning system for Parkinsonian gait analysis”. En: *International Conference on Biomedical and Health Informatics* (2018), págs. 108-111 (vid. pág. 12).

BELIC, Minja. “Artificial intelligence for assisting diagnostics and assessment of Parkinson’s disease – A review”. En: *Clinical neurology and neurosurgery* (2019), pág. 105442 (vid. pág. 11).

CARAMIA, Carlotta. “IMU-Based Classification of Parkinson’s Disease From Gait: A Sensitivity Analysis on Sensor Location and Feature Selection”. En: *IEEE journal of biomedical and health informatics* 22.6 (2018), págs. 1765-1774 (vid. pág. 11).

CHALAPATHY, R., CHAWLA, S.. “Deep learning for anomaly detection: A survey”. En: (2019) (vid. pág. 24).

CHANDOLA, V., BANERJEE, A. y KUMAR, V.. “Outlier detection: A survey.” En: *ACM Computing Surveys* (2007) (vid. págs. 24, 25).

DI MATTIA, Federico GALEONE, Paolo, DE SIMONI, Michelle GHELFI, Emanuel. “A Survey on GANs for Anomaly Detection”. En: (2019) (vid. págs. 25, 28).

EKKER, Merel, JANSSEN, Sabine y SEPPI, Klaus. “Ocular and visual disorders in Parkinson’s disease: Common but frequently overlooked”. En: *Parkinsonism and related disorders* 40 (2017), págs. 1-10 (vid. pág. 13).

GITCHEL, George. “Experimental support that ocular tremor in Parkinson’s disease does not originate from head movement”. En: *Parkinsonism and related disorders* 20.7 (2014), 743–747 (vid. pág. 34).

GITCHEL, George, WETZEL, Paul y BARON, Mark. “Pervasive ocular tremor in patients with Parkinson disease”. En: *Archives of neurology* 69.8 (2012), págs. 1011-1017 (vid. págs. 11, 14, 34).

GOLDBERG, M.E. y WALKER, M.F.. “The control of gaze”. En: *Principles of Neural Science*. (2013), 894–916 (vid. pág. 14).

GOODFELLOW, Ian POUGET-ABADIE, Jean, MIRZA, Mehdi XU, Bing WARDE-FARLEY, David OZAIR, Serjhil COURVILLE, Aaron y BENGIO, Yoshua. “Generative adversarial nets”. En: *Advances in Neural Information Processing Systems* (2014), 2672–2680 (vid. pág. 18).

GORGES, Martin, MÜLLER, Peter-Hans y LULÉ, DOROTHEÉ. “The association between alterations of eye movement control and cerebral intrinsic functional connectivity in Parkinson’s disease”. En: *Brain imaging and behavior* 10.1 (2016), págs. 79-91 (vid. pág. 13).

HAWKINS, D. *Identification of Outliers*. Chapman y Hall, 1980 (vid. pág. 24).

HU, Kun. “Vision-based freezing of gait detection with anatomic patch based representation”. En: *Asian Conference on Computer Vision*. (2018), págs. 564-576 (vid. pág. 12).

IRANZO, Alex, SANTAMARIA, Joan, y TOLOSA, Eduardo. “Idiopathic rapid eye movement sleep behaviour disorder: diagnosis, management, and the need for neuroprotective interventions”. En: *The Lancet Neurology* 15.4 (2016), págs. 405-419 (vid. pág. 13).

JANKOVIC, Joseph. “Parkinson’s disease: clinical features and diagnosis”. En: *Journal of neurology, neurosurgery and psychiatry* 79.4 (2008), págs. 368-376 (vid. pág. 10).

JEPSON, Allan D., FLEET, David J.and. “Computation of component image velocity from local phase information”. En: *International Journal of computer vision* 5.1 (1990), 77–104 (vid. pág. 34).

KASKI, D. SAIFEE, T.A., BUCKWELL, D.. “Ocular tremor in Parkinson’s Disease Is due to head oscillation”. En: *Movement Disorders* 28 (2013) (vid. págs. 11, 14, 24).

KHOSLA, Ajit y KIM, Dongsoo. *Optical Imaging Devices: New Technologies and Applications*. CRC Press, 2015 (vid. pág. 16).

KUBIS, Anna, SZYMANSKI, Artur y PRZYBYSZEWSKI, Andrezej W.. “Fuzzy Rough sets theory applied to parameters of eye movements can help to predict effects of different treatments in Parkinson’s patients”. En: *International Conference on Pattern Recognition and Machine Intelligence* (2015), págs. 325-334 (vid. pág. 16).

LAI, Hsin-Yu. “Enabling Saccade Latency Measurements with Consumer-Grade Cameras”. En: *International Conference on Image Processing (ICIP)* 25 (2018), 3169–3173 (vid. pág. 16).

LARRAZABAL, A.J., GARCÍA, C.E. y MARTÍNEZ, F.. “Video-oculography eye tracking towards clinical applications: A review”. En: *Computers in biology and medicine* 108 (2019), 57–66 (vid. pág. 14).

MIRZA, M. y OSINDERO, S. “Conditional Generative Adversarial Nets”. En: (2014) (vid. pág. 22).

PAFFENROTH, R., KAY, K. y SERVI, L.. “Robust PCA for Anomaly Detection in CyberNetworks”. En: (2018) (vid. pág. 25).

POEWE, Werner. “Parkinson’s disease”. En: *Nature reviews Disease primers 3* (2017), pág. 17013 (vid. págs. 10, 11).

PORTILLA, Javier y SIMONCELLI, Eero. “A parametric texture model based on joint statistics of complex wavelet coefficients”. En: *Internacional Journal of computer vision* 40.1 (2000), 49–70 (vid. pág. 35).

RIZZO, Giovanni, COPETTI, Massimiliano, y ARCUTI, Simona. “Accuracy of clinical diagnosis of Parkinson disease A systematic review and meta-analysis”. En: *Neurology* 86.6 (2016), págs. 566-576 (vid. pág. 10).

SALAT, David. “Challenges of modifying disease progression in pre-diagnostic Parkinson’s disease”. En: *The Lancet Neurology* 15.6 (2016), págs. 637-648.7 (vid. pág. 12).

SALAZAR, I, PERTUZ, S.D. y MARTINEZ, F. “Visual amplification of oculomotor signs for parkinsonian patterns recognition”. En: *Computers in biology and medicine* (2020), 8–10 (vid. págs. 12, 14, 16, 17).

SCHLEGL, T., SEEBCK, P., WALDSTEIN, S.M. SCHMIDT-ERFURTH, U. y LANGS, G.. “Unsupervised Anomaly Detection with Generative Adversarial Networks to Guide Marker Discovery”. En: (2017) (vid. pág. 28).

- SLEDZIANOWSKI, Albert. “Measurements of Antisaccades Parameters Can Improve the Prediction of Parkinson’s Disease Progression”. En: *Asian Conference on Intelligent Information and Database Systems* (2019), 602–614 (vid. pág. 15).
- SZYMANSKI, Artur. “Building Intelligent Classifiers for Doctor-Independent Parkinson’s Disease Treatments”. En: *Conference of Information Technologies in Biomedicine* (2016), págs. 267-276 (vid. pág. 15).
- TRUJILLO, David MARTÍNEZ, Fabio, ATEHORTÚAL, Angélica. “A characterization of Parkinson’s disease by describing the visual field motion during gait”. En: *11th International Symposium on Medical Information Processing and Analysis* 9681 (2015) (vid. pág. 16).
- TURCANO, Pierpaolo. “Early ophthalmologic features of Parkinson’s disease: a review of preceding clinical and diagnostic markers”. En: *Journal of neurology* (2018), págs. 1-9 (vid. pág. 13).
- VENUTO, Charles. “A review of disease progression models of Parkinson’s disease and applications in clinical trials”. En: *Movement Disorders* 31.4 (2016), págs. 368-376 (vid. pág. 10).
- ZENATI, H. FOO, C.S., LECOAT, B. MANEK, G. y CHANDRASEKHAR. V.R.. “Efficient GAN-Based Anomaly Detection”. En: (2018) (vid. págs. 23, 28).
- ZHANG, Yichao, PINTEA, Silvia y VAN GEMERT, Jan. “Video Acceleration Magnification”. En: *Computer Vision and Pattern Recognition* (2017) (vid. pág. 34).
- ZHANG, Yu. “Oculomotor Performances Are Associated With Motor and Non-motor Symptoms in Parkinson’s Disease”. En: *Frontiers in neurology* 9 (2018), pág. 960 (vid. pág. 15).
- ZHOU, C y PAFFENROTH, R. “Anomaly Detection with Robust Deep Autoencoders”. En: (2017) (vid. pág. 27).