



**REPRESENTACIONES AUTO-SUPERVISADAS DE MATRICES SPD PARA
LA CUANTIFICACIÓN DE PATRONES DE FIJACIÓN OCULAR EN
PACIENTES CON PARKINSON**

EDWARD ANDRES SANDOVAL PINEDA

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

2023

**REPRESENTACIONES AUTO-SUPERVISADAS DE MATRICES SPD PARA
LA CUANTIFICACIÓN DE PATRONES DE FIJACIÓN OCULAR EN
PACIENTES CON PARKINSON**

EDWARD ANDRES SANDOVAL PINEDA

**Trabajo de Grado para optar por el título de:
Ingeniero de Sistemas**

Director:

Fabio Martínez Carrillo

Ph.D in Systems and Computer Engineering

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

2023

AGRADECIMIENTOS

El autor expresa sus agradecimientos:

Primero que todo, al Profesor Fabio Martínez, director del grupo de investigación BIVL2ab por su guía durante el desarrollo de este proyecto. Su experiencia y conocimientos en el campo de la inteligencia artificial fueron fundamentales para el éxito de mi proyecto. También quiero agradecer a Juan Olmos, matemático que brindó un apoyo constante y valioso en el desarrollo del proyecto. Su experiencia en el área de trabajo y trayectoria académica fueron clave para llevar a cabo la implementación de mi trabajo.

Quiero agradecer a mi familia por el apoyo incondicional que me han brindado durante toda mi carrera y en especial durante la elaboración de este proyecto. Su confianza en mis capacidades ha sido fundamental para alcanzar los objetivos que me he propuesto. También quiero a mi pareja por el apoyo y motivación que me ha brindado en momentos clave de este proceso. Su paciencia y comprensión han sido de suma importancia para lograr culminar esta etapa de mi formación académica. Agradezco a mis amigos por haberme acompañado y alentado durante todo este tiempo. Su ánimo y disposición a escuchar mis ideas y reflexiones han sido de gran ayuda para mi crecimiento personal y profesional.

También, quiero expresar mi agradecimiento a la Escuela de Ingeniería de Sistemas por brindarme una formación de calidad y completa durante estos años de estudio. Los conocimientos adquiridos en esta institución me han permitido enfrentar con éxito los desafíos que se me han presentado a lo largo de mi proyecto.

Por último, agradecer al anillo de nucita que estuvo ahí de intruso mientras terminaba de escribir.

CONTENIDO

	pág
INTRODUCCIÓN	10
1. FUNDAMENTOS Y TRABAJO PREVIO	14
1.1. La enfermedad de Parkinson y los patrones oculomotores	14
1.1.1. Fijación Ocular.	15
1.1.2. Movimiento suave.	15
1.2. Representaciones de aprendizaje y clasificación	16
1.2.1. Autoencoders	17
1.3. Descriptores SPD	19
1.4. Caracterización oculomotora del Parkinson	21
2. PROBLEMA DE INVESTIGACIÓN	23
3. OBJETIVOS	25
4. Conjunto de Datos	26
5. Método Propuesto	28
5.1. Descriptor de patrones de fijación	28
5.2. Autoencoder SPD	30
5.2.1. Codificador de matrices SPD	30
5.2.2. Decodificador de matrices SPD	32
5.2.3. Esquema de aprendizaje	34
5.3. Clasificación de embebidos	38
5.4. Configuración experimental	40

5.4.1. Mapas de activación 40

5.4.2. Arquitectura codificador-decodificador 41

5.4.3. Validación 42

5.4.4. Clasificación de embebidos 42

6. EVALUACIÓN Y RESULTADOS 43

7. CONCLUSIONES Y TRABAJO FUTURO 50

BIBLIOGRAFIA 52

LISTA DE FIGURAS

	pág
Figura 1. Figura Autoencoder	18
Figura 2. Mapeos Riemannianos y SPD-Net	20
Figura 3. Cortes temporales	27
Figura 4. Pipeline del método propuesto	29
Figura 5. Descriptor SPD	30
Figura 6. Embebido 3d de los dos mejores métodos	47
Figura 7. Curvas ROC de los dos mejores métodos	48

LISTA DE TABLAS

	pág
Tabla 1. Resultados Errores de Reconstrucción con ajuste L_{MSE}	44
Tabla 2. Resultados Errores de Reconstrucción $L_{MSE-log}$	44
Tabla 3. Clasificación de embebidos ajustados con datos VGG19	46
Tabla 4. Clasificación de embebidos ajustados con datos Xception	46
Tabla 5. Comparativa con métodos del estado del arte	49

RESUMEN

TÍTULO: REPRESENTACIONES AUTO-SUPERVISADAS DE MATRICES SPD PARA LA CUANTIFICACIÓN DE PATRONES DE FIJACIÓN OCULAR EN PACIENTES CON PARKINSON *

AUTOR: EDWARD ANDRES SANDOVAL PINEDA **

PALABRAS CLAVE: MATRICES SPD, APRENDIZAJE RIEMANNIANO, REPRESENTACIÓN AUTO-SUPERVISADA, ENFERMEDAD DE PARKINSON, PATRONES OCULOMOTORES

DESCRIPCIÓN: La enfermedad de Parkinson (EP) es el segundo trastorno neurodegenerativo más común en el mundo, afectando entre el 2 y 3 % de la población mayor a 65 años. Esta enfermedad no tiene cura y hasta ahora no se conoce un biomarcador definitivo para la detección y cuantificación de la enfermedad. Recientemente las alteraciones oculomotoras durante tareas de fijación ocular han demostrado evidencias prometedoras como descriptor de la enfermedad, con suficiente sensibilidad para discriminar entre estadios de la enfermedad, incluso en etapas tempranas. No obstante, los métodos computacionales para soportar el diagnóstico y apoyar la cuantificación son basados en procesos estrictamente supervisados, requiriendo grandes volúmenes de datos e incluyendo un sesgo asociado a la subjetividad de las observaciones de expertos. En este trabajo se propone una arquitectura auto-supervisada de caracter Riemanniano que puede aprender patrones de fijación oculomotor sin supervisión. Los descriptores embebidos resultantes de esta representación tiene la capacidad de discriminar secuencias de video de personas afectadas con la EP con respecto a una población control. Las observaciones de fijación son proyectadas a una red convolucional pre-entrenada para obtener una descripción en términos de características profundas de nivel intermedio. Estas características son a su vez codificadas en matrices simétricas definidas positivas (SPD), que sirven como entrada a la representación la arquitectura auto-supervisada. Entonces, la representación propuesta es capaz de aprovechar la estructura geométrica de los datos, considerando al variedad de Riemann, para obtener descriptores embebidos geométricos y compactos. Estos vectores en una tarea de clasificación, en un estudio con 13 pacientes control y 13 pacientes diagnosticados con la EP, logró una exactitud de 92,3 % y un AUC de 0,97 .

* Trabajo de investigación

** FFacultad de Ingenierías Físicomecánicas. Escuela de Ingeniería de Sistemas e Informática. Director: Fabio Martínez, PhD.

ABSTRACT

TITLE: SELF-SUPERVISED REPRESENTATIONS OF SPD MATRICES FOR THE QUANTIFICATION OF OCULAR FIXATION PATTERNS IN PARKINSON PATIENTS *

AUTHOR: EDWARD ANDRES SANDOVAL PINEDA **

KEYWORDS: SPD MATRICES, RIEMANNIAN LEARNING, SELF-SUPERVISED REPRESENTATION, PARKINSON'S DISEASE, OCULOMOTOR PATTERNS.

DESCRIPTION: Parkinson's disease (PD) is the second most common neurodegenerative disorder in the world, affecting between 2 and 3% of the population over 65 years old. This disease has no cure, and so far, a definitive biomarker for the detection and quantification of the disease is not known. Recently, oculomotor alterations during eye fixation tasks have shown promising evidence as a disease descriptor, with sufficient sensitivity to discriminate between disease stages, even in the early stages. However, computational methods to support diagnosis and aid in quantification are based on strictly supervised processes, requiring large volumes of data and including a bias associated with the subjectivity of expert observations. In this work, a self-supervised architecture based on Riemannian features is proposed, which can learn oculomotor fixation patterns without supervision. The resulting embedded descriptors from this representation have the ability to discriminate video sequences of individuals affected by PD compared to a control population. The fixation observations are projected onto a pre-trained convolutional network to obtain a description in terms of deep intermediate features. These features are then encoded into symmetric positive definite (SPD) matrices, which serve as input to the self-supervised architecture representation. Thus, the proposed representation is capable of leveraging the geometric structure of the data, considering the Riemannian manifold, to obtain geometric and compact embedded descriptors. During a classification task, in a study with 13 control patients and 13 patients diagnosed with PD, these embeddeds achieved an accuracy of 92.3% and an AUC of 0.97.

* Research work

** Faculty of Physics-Mechanics Engineering. School of Systems Engineering and Informatics. Advisor: Fabio Martínez, PhD.

INTRODUCCIÓN

Las enfermedades neurológicas son la principal causa de discapacidad en el mundo, siendo el Parkinson, la enfermedad que más rápido ha incrementado en prevalencia, discapacidad y muertes, afectando entre un 2-3 % de la población mundial mayor a 65 años para 2019 ¹. Actualmente la enfermedad del Parkinson (EP) no tiene cura, sin embargo, la planificación de un tratamiento personalizado desde etapas tempranas es decisiva para mejorar su pronóstico y reducir la probabilidad de discapacidades motoras ¹.

La EP es consecuencia directa de un descenso gradual de los niveles de dopamina, un neurotransmisor del sistema nervioso central que regula las órdenes motoras ². Entre los síntomas motores mas comunes se puede mencionar el temblor, la inestabilidad postural y la lentitud en movimientos ³. Sin embargo, estos síntomas se presentan comúnmente en etapas avanzadas de la enfermedad, lo cual dificulta el diagnóstico oportuno y temprano de la enfermedad. Además, el diagnóstico se basa principalmente en el análisis observacional de patrones motores, el cual es afectado por errores de subjetividad del profesional. De hecho, se ha reportado que mas de un 24 % de los diagnósticos son pasados por alto en etapas tempranas ⁴.

Recientemente, los patrones oculomotores durante tareas de fijación ocular han demostrado ser

¹ Valery L. Feigin, Emma Nichols, Tahiya Alam et al. “Global, regional, and national burden of neurological disorders, 1990–2016: a systematic analysis for the Global Burden of Disease Study 2016”. En: *The Lancet Neurology* 18.5 (2019), págs. 459-480.

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

³ Rimona S. Weil, Anette E. Schrag, Jason D. Warren et al. “Visual dysfunction in Parkinson’s disease”. En: *Brain* 139.11 (2016), págs. 2827-2843.

⁴ Werner Poewe, Klaus Seppi, Caroline M. Tanner et al. “Parkinson disease”. En: *Nature Reviews Disease Primers* 3 (2017), págs. 1-21.

un potencial biomarcador de la EP desde etapas tempranas ^{5 6 7}. Tomando esto en consideración, se han propuesto métodos para el reconocimiento de anomalías, asociadas a la EP, en dichos patrones oculomotores. Sin embargo, estos métodos incluyen configuraciones experimentales que requieren equipos sofisticados o dispositivos que solo consideran trayectorias generales del movimiento ocular ^{8 3 9}. Como alternativa se han propuesto métodos computacionales para el análisis de patrones motores directamente de secuencias de vídeo ^{10 11 12 13}. Entre estos, los enfoques de aprendizaje profundo han revelado ventajas determinantes para descubrir patrones ocultos y caracterizar descriptores cinemáticos en la modelación de trastornos del movimiento ocular, el seguimiento del movimiento ocular, entre otros ^{12 13}. Dada la variabilidad de los datos,

-
- ⁵ Merel S Ekker, Sabine Janssen, Klaus Seppi et al. “Ocular and visual disorders in Parkinson’s disease: common but frequently overlooked”. En: *Parkinsonism & related disorders* 40 (2017), págs. 1-10.
- ⁶ Pierpaolo Turcano et al. “Early ophthalmologic features of Parkinson’s disease: a review of preceding clinical and diagnostic markers”. En: *Journal of neurology* 266 (2019), págs. 2103-2111.
- ⁷ Vivek Lal y Daniel Truong. “Eye movement abnormalities in movement disorders”. En: *Clinical Parkinsonism & Related Disorders* 1 (2019), págs. 54-63.
- ⁸ Agostina J Larrazabal, CE García Cena y César Ernesto Martínez. “Video-oculography eye tracking towards clinical applications: A review”. En: *Computers in biology and medicine* 108 (2019), págs. 57-66.
- ⁹ George T Gitchel, Paul A Wetzel y Mark S Baron. “Pervasive ocular tremor in patients with Parkinson disease”. En: *Archives of neurology* 69.8 (2012), págs. 1011-1017.
- ¹⁰ Hsin-Yu Lai, Gladynel Saavedra-Peña, Charles Sodini et al. “Enabling Saccade Latency Measurements with Consumer-Grade Cameras”. En: *2018 25th IEEE International Conference on Image Processing (ICIP)*. 2018, págs. 3169-3173.
- ¹¹ Tana B Carson y Stephen Z Sutton. *Application for smart phone or related devices for use in assessment of vestibulo-ocular reflex function*. US Patent App. 15/569,472. 2018.
- ¹² Wu Wang et al. “Early detection of Parkinson’s disease using deep learning and machine learning”. En: *IEEE Access* 8 (2020), págs. 147635-147646.
- ¹³ Andrzej W Przybyszewski, Stanislaw Szlufik et al. “Machine learning on the video basis of slow pursuit eye movements can predict symptom development in Parkinson’s patients”. En: *Intelligent Information and Database Systems: 7th Asian Conference, ACIIDS 2015, Bali, Indonesia, March 23-25, 2015, Proceedings, Part II* 7. Springer International Publishing. 2015, págs. 268-276.

estos métodos requieren de grandes conjuntos de entrenamiento y anotaciones de expertos para obtener generalizaciones adecuadas de los mismos ¹². Este desafío ha sido abordado desde representaciones generales y pre-entrenadas para obtener un banco de descriptores. Posteriormente se utilizan descriptores de baja dimensionalidad para compactar esta información ^{14 15 16}. Estos descriptores permiten guardar relaciones estadísticas que requieren un estudio cuidadoso de las restricciones geométricas. En este sentido, recientes modelos han considerado el modelamiento de la geometría de los datos usando representaciones compactas ^{16 17}. Estos modelos han logrado afrontar la escasez de la información para el entrenamiento, no obstante, estos métodos siguen siendo dependiente de las anotaciones brindadas por los expertos. En este sentido, estas representaciones discriminativas y supervisadas pueden estar sesgadas por la información brindada la calidad de las etiquetas.

Este trabajo propone una arquitectura autosupervisada y compacta que permite modelar patrones asociados a la fijación ocular. Los vectores embebidos resultantes de la representación pueden ser usados en una tarea de discriminación, logrando separar ejemplos de secuencias de una población control con respecto a una población diagnosticada con la EP. Para esto, cada video de fijación ocular es transformado en un corte 2d espacio-temporal que captura patrones temporales asociados al temblor ocular durante la tarea de fijación. Posteriormente, cada corte se describe por mapas de activación convolucionales que posteriormente se codifican en una matriz compacta SPD que guarda relaciones e información relevantes entre las características

¹⁴ Isail Salazar et al. “A convolutional oculomotor representation to model parkinsonian fixational patterns from magnified videos”. En: *Pattern Analysis and Applications* 24 (2021), págs. 445-457.

¹⁵ Dinesh Acharya et al. “Covariance pooling for facial expression recognition”. En: *Proceedings of the IEEE conference on computer vision and pattern recognition workshops*. 2018, págs. 367-374.

¹⁶ Juan Olmos, Antoine Manzanera y Fabio Martínez. “An Oculomotor Digital Parkinson Biomarker from a Deep Riemannian Representation”. En: *Pattern Recognition and Artificial Intelligence: Third International Conference, ICPRAI 2022, Paris, France, June 1–3, 2022, Proceedings, Part I*. Springer. 2022, págs. 677-687.

¹⁷ Zhiwu Huang y Luc Van Gool. “A riemannian network for spd matrix learning”. En: *Thirty-first AAAI conference on artificial intelligence*. 2017.

profundas. A partir de esta entrada SPD, se propone un esquema de auto-aprendizaje de matrices SPD por medio de una arquitectura codificador-decodificador que permite el ajuste de representaciones embebidas, sin el requerimiento de etiquetas relacionadas con la enfermedad, por medio de una función de pérdida de reconstrucción que respeta la geometría de los datos. Para evaluar el trabajo propuesto se seleccionó un conjunto de datos de pacientes diagnosticados con la EP y una población control. El esquema fue evaluado por su capacidad de reconstrucción, y además se comparó el método propuesto con respecto a estrategias de aprendizaje supervisadas midiendo la capacidad de clasificar representaciones profundas embebidas provenientes de una población Parkinson y control.

1. FUNDAMENTOS Y TRABAJO PREVIO

1.1. La enfermedad de Parkinson y los patrones oculomotores

La enfermedad de Parkinson (EP) se caracteriza por la pérdida de neurotransmisores, producto del déficit de dopamina¹⁸. Esta pérdida afecta al sistema nervioso, provocando alteraciones en el movimiento. Algunos de los síntomas más comunes para cuantificar anomalías relacionadas con la enfermedad incluyen el temblor, la inestabilidad en la postura y la lentitud en movimientos⁴. Sin embargo, estos síntomas se hacen evidentes en estadios tardíos de la enfermedad y son evaluados con una alta subjetividad, alcanzando un error de diagnóstico por encima de 24%⁴. Actualmente, no existe cura para la enfermedad y tampoco hay un biomarcador definitivo de la enfermedad. Por tanto, el tratamiento oportuno y metodologías para el diagnóstico en etapas tempranas es fundamental para retardar la aparición de síntomas más severos que provoquen disfuncionalidades.

En los últimos años, los patrones oculomotores han emergido como un potencial biomarcador temprano de la enfermedad¹⁹. Varios trabajos encontraron experimentalmente correlaciones entre las alteraciones en el control del movimiento ocular y la EP^{20 21}. También en la literatura se han reportado evidencias sobre una correlación entre cambios neurodegenerativos por el avance de la enfermedad y alteraciones de patrones oculomotores específicos^{5 6 7}. Dentro de

¹⁸ JW Błaszczyk. "Motor deficiency in Parkinson's disease". En: *Acta neurobiologiae experimentalis* 58.1 (1998), 79—93.

¹⁹ CA Antoniades y C Kennard. "Ocular motor abnormalities in neurodegenerative disorders". En: *Eye* 29.2 (2015), pág. 200.

²⁰ Cindy Ezdon. "Principles of Neural Science". En: *The American Journal of Occupational Therapy* 41.1 (1987), págs. 59-59.

²¹ Michael R. MacAskill y Tim J. Anderson. "Eye movements in neurodegenerative diseases". En: *Current Opinion in Neurology* 29.1 (2016).

las diferentes anormalidades que han sido observadas en estudios, las alteraciones en tareas de fijación ocular y seguimiento suave son las más comunes. A continuación se describe brevemente cada una de estas tareas.

1.1.1. Fijación Ocular. En esta tarea los participantes deben estabilizar la vista en un punto fijo, mientras la cabeza se mantiene estática. Durante la tarea de fijación, se pueden presentar varios tipos de alteraciones del movimiento; como por ejemplo la deriva ocular (un movimiento sinuoso y lento que produce cambios mayores en la posición del ojo); microsacadas, movimientos más amplios y rápidos con frecuencias de 1-2Hz y amplitud $< 0, 2^\circ$; micro-temblores, movimientos oscilatorios de alta frecuencia (40-100Hz) y poca amplitud ($\sim 0, 002^\circ$)²². Se ha comprobado empíricamente que en el ejercicio de fijación ocular persisten estas alteraciones y combinaciones de ellas en los pacientes con la EP^{23 9 3}. Por ejemplo, en un estudio realizado por Gitchel et al con 112 pacientes (60 control y 52 con EP)⁹ se encontró una anomalía de inestabilidad presente en todos los pacientes con la EP de una población de estudio, concretamente una oscilación con 5.7 Hz de frecuencia y 0.27° de amplitud en promedio, mientras que en los sujetos control 58 pacientes mostraron un comportamiento estable, no oscilatorio. Debido a tu alta sensibilidad demostrada, en el presente trabajo se propondrá un método que a partir de estas anomalías de fijación pueda aprender representaciones discriminativas y patrones oculomotores relacionados con la EP.

1.1.2. Movimiento suave. En esta tarea, los participantes deben realizar un seguimiento visual de la trayectoria de un estímulo ubicado en frente de ellos que realiza movimientos horizontales o verticales a una velocidad constante. En este tipo de tareas de movimiento ocular

²² Richard J Krauzlis, Laurent Goffart y Ziad M Hafed. “Neuronal control of fixation and fixational eye movements”. En: *Philosophical Transactions of the Royal Society B: Biological Sciences* 372.1718 (2017).

²³ Christian Duval y Anne Beuter. “Fluctuations in Tremor at Rest and Eye Movements during Ocular Fixation in Subjects with Parkinson’s Disease”. En: *Parkinsonism & Related Disorders* 4 (1998), págs. 91-97.

existe evidencia que sugiere una relación entre anomalías del seguimiento ocular y el deterioro de la EP ²⁴ ²⁵ ²⁶. Por ejemplo, Fukushima et al ²⁴ encontró que más del 75% de los pacientes con EP presentan un retraso en iniciar de forma sincronizada el seguimiento suave del objeto. Además, durante el movimiento se observaron interrupciones por movimientos sacádicos que surgían al intentar corregir la trayectoria de la mirada hacia la trayectoria del estímulo. Al igual que la fijación, esta tarea oculomotora constituye un potencial biomarcador de la EP ²⁶. Sin embargo, capturar la dinámica del movimiento ocular para encontrar anomalías constituye un reto complejo que se incrementa teniendo en cuenta la limitación de la cantidad de datos.

1.2. Representaciones de aprendizaje y clasificación

Los algoritmos de aprendizaje para la clasificación de patrones anormales son esquemas típicamente supervisados, que permiten construir modelos robustos de discriminación, usando un conjunto de datos de muestras etiquetadas $\{(x_i, y_i)\}_{i=1}^N$, donde cada x_i denota las características de entrada y y_i su correspondiente etiqueta ²⁷. Entonces, el modelo es ajustado al encontrar los parámetros del modelo θ que predicen mejor los datos minimizando una función de pérdida $L(y, \hat{y})$ donde $\hat{y} = f(X; \theta)$ ²⁸. Actualmente, el esquema más frecuente de aprendizaje se basa

²⁴ Kikuro Fukushima, Norie Ito, Graham R Barnes et al. “Impaired smooth-pursuit in Parkinson’s disease: normal cue-information memory, but dysfunction of extra-retinal mechanisms for pursuit preparation and execution”. En: *Physiological reports* 3.3 (2015), e12361.

²⁵ Hiroshi Shibasaki, Sadatoshi Tsuji y Yoshigoro Kuroiwa. “Oculomotor Abnormalities in Parkinson’s Disease”. En: *Archives of Neurology* 36.6 (jun. de 1979), págs. 360-364.

²⁶ R. A. Armstrong. “Oculo-Visual Dysfunction in Parkinson’s Disease”. En: *Journal of Parkinson’s Disease* 5 (2015). 4, págs. 715-726.

²⁷ Vijay K. Vemuri. “The Hundred-Page Machine Learning Book”. En: *Journal of Information Technology Case and Application Research* 22.2 (2020), págs. 136-138.

²⁸ Geert Litjens, Thijs Kooi, Babak Ehteshami Bejnordi et al. “A survey on deep learning in medical image analysis”. En: *Medical Image Analysis* 42 (2017), págs. 60-88.

en redes neuronales profundas que constan de un conjunto de unidades no lineales (neuronas), agrupadas en diferentes niveles de procesamiento, formando el modelo como una función que consiste en una sucesión de capas (ecuación 1) y es ajustable por medio de algoritmos iterativos como el gradiente descendiente.

$$f(x, \theta) = \sigma_L(W_{L-1} * \sigma_{L-1}(W_{L-2} * \cdots * \sigma_1(W_1 * x + b)) + b). \quad (1)$$

Los operadores $*$ más comunes son combinaciones lineales, para las redes neuronales profundas, y convoluciones discretas, para las redes neuronales convolucionales. Sin embargo, estos modelos están sujetos a las anotaciones, con lo cual pueden sufrir sesgos y tendencias de sobreaprendizaje, que dificultan la generalización de la representación para su respectiva tarea ^{29 30}. Por tanto, nuevas arquitecturas y estrategias de aprendizaje intentan involucrar metodologías autosupervisadas, semi-supervisadas y no-supervisadas para mejorar el comportamiento de los modelos, aprovechando información no etiquetada disponible.

1.2.1. Autoencoders Dentro de las estrategias no-supervisadas, sobresalen las estrategias de autoaprendizaje (autoencoders) que tienen como objetivo aprender representaciones de baja dimensionalidad. Para esto, a partir del dato de entrada X , estas estrategias utilizan un módulo codificador E_Θ y decodificador D_{Θ_a} para generar una representación embebida de baja dimensión $Z = E_\Theta(X)$ que posteriormente es utilizada para generar una reconstrucción $\hat{X} = D_\Phi(Z)$ (Ver Figura 1). El aprendizaje de estos modelos busca unos pesos óptimos Θ, Φ que minimicen una

²⁹ Zhi-Hua Zhou. “A brief introduction to weakly supervised learning”. En: *National science review* 5.1 (2018), págs. 44-53.

³⁰ Xue Ying. “An overview of overfitting and its solutions”. En: *Journal of physics: Conference series*. Vol. 1168. IOP Publishing. 2019, pág. 022022.

función de pérdida de reconstrucción $L(X, \hat{X})$ ³¹.

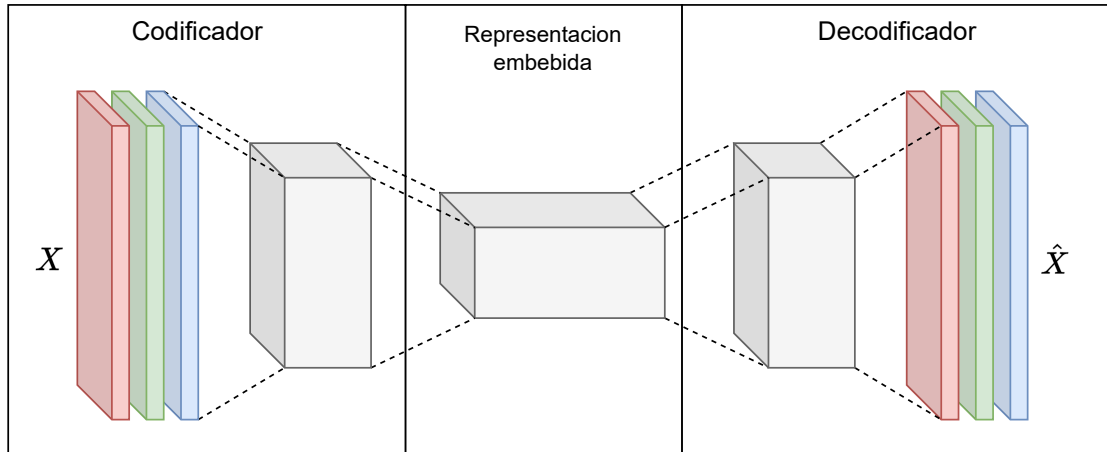


Figura 1. Red estilo Codificador-Decodificador (Autoencoder). Por medio de una tarea de pretexto que puede ser autosupervisada (reconstrucción), busca generar representaciones embebidas de menor dimensión

Dentro de las ventajas de estos modelos, está su capacidad para extraer características significativas relacionadas con el problema al cual fueron ajustados. Además, aprende a generar un espacio embebido de baja dimensionalidad donde los datos son proyectados, permitiendo la caracterización o discriminación entre los datos del problema. Sin embargo, estas estrategias de aprendizaje siguen siendo dependientes de una gran cantidad de datos para lidiar con la alta variabilidad. Por el contrario, una poca cantidad de datos no permite alcanzar una representación discriminativa representativa en el espacio latente de los embebidos. Para lidiar con esto, en la literatura se han propuesto descriptores compactos que reducen la dimensión mientras guardan medidas estadísticas como la covarianza u otras estimaciones de similitud.

³¹ Pierre Baldi. “Autoencoders, unsupervised learning, and deep architectures”. En: *Proceedings of ICML workshop on unsupervised and transfer learning*. JMLR Workshop y Conference Proceedings. 2012, págs. 37-49.

1.3. Descriptores SPD

Las matrices de covarianza y correlación son ejemplos, de matrices simétricas definidas positivas (SPD), que han sido propuestos para describir relaciones lineales y de segundo orden ³². Este tipo de descriptores han sido implementados en diferentes áreas de la visión por computador para compactar representaciones de alta dimensión y capturar información de segundo orden relevantes. Para esto, dado un conjunto de vectores $X = \{x_i\}_{i=1}^m$, la matriz de covarianza se define como $C = [\gamma_{ij}]_{i=1,j=1}^{m,m}$, donde $\gamma_{ij} = E[(x_i - E[x_i])(y_i - E[y_i])]$ es la covarianza entre los vectores x_i y x_j . De esta forma, la matriz de covarianza guarda en sus entradas medidas de similitud entre todos los distintos pares de características. De forma similar, se pueden definir matrices SPD a partir de medidas de similitud γ_{ij} diferentes.

Estas matrices simétricas definidas positivas, se denotan como matrices que pertenecen a Sym_d^+ , donde d es la dimensión de la matriz. Además, forman un espacio geométrico en una variedad Riemanniana, la cual, es un cono convexo diferenciable de curvatura positiva (no plano). Por tanto, operar sobre estas matrices requiere operaciones apropiadas que tengan en cuenta las propiedades geométricas de este espacio. Una aproximación para esto ha sido utilizar transformaciones desde y hacia un espacio Euclidiano no curvo ³³. Por ejemplo, se ha utilizado el plano tangente, el cual es un espacio Euclidiano tangente a Sym_d^+ . Para esto, al cual se pueden proyectar los datos desde la variedad SPD por medio de la función logaritmo. Para esto, dada una matriz $A \in Sym_d^+$, esta se puede descomponer en autovalores y autovectores $A = ZDZ^T$, donde $D = \text{diag}(\lambda_i)$ es una matriz diagonal cuyas entradas son los autovalores asociados a A y las columnas de Z son los respectivos autovectores. Se calcula entonces $f(A)$ para hallar su proyección, donde f es la función logaritmo, de la siguiente forma

³² Hà Quang Minh y Vittorio Murino. *Covariances in computer vision and machine learning*. Springer, 2018, págs. 7-17.

³³ Oncel Tuzel, Fatih Murat Porikli y Peter Meer. "Pedestrian Detection via Classification on Riemannian Manifolds". En: *IEEE Transactions on Pattern Analysis and Machine Intelligence* 30 (2008), págs. 1713-1727.

$$f(A) = Zf(D)Z^T = Zf(\text{diag}(\lambda_i))Z^T. \quad (2)$$

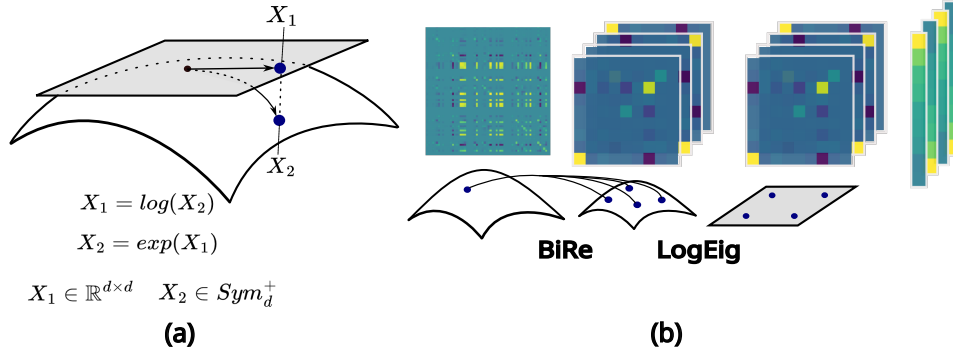


Figura 2. (a) Ilustración de los mapeos con el operador \log o \exp que proyectan un dato desde el espacio Riemanniano al espacio euclidiano y viceversa. (b) Se observa el funcionamiento de la red para aprendizaje de representaciones que conserven propiedades geométricas SPD-Net.

De la misma manera, se puede realizar la proyección inversa, desde el plano tangente hacia el espacio Riemanniano, para trasportar de nuevo los datos a Sym_d^+ . Para esto, se realiza el mismo cálculo que en (2), tomando ahora a f como la función exponencial. Una ilustración de estas funciones se presenta en la Figura 2 a.

En la literatura también se han logrado integrar representaciones SPD con esquemas actuales de aprendizaje, permitiendo aprovechar la geometría de los datos para construir representaciones que se ajusten apropiadamente a una regla de minimización. Recientemente, se han hecho esfuerzos por proponer métodos que preserven la geometría de los datos y puedan minimizar tareas de aprendizaje de forma más eficiente a partir de datos en espacios no-Euclidianos ³⁴

³⁴ Vincent Arsigny et al. “Geometric Means in a Novel Vector Space Structure on Symmetric Positive-Definite Matrices”. En: *SIAM Journal on Matrix Analysis and Applications* 29.1 (2007), págs. 328-347.

³⁵ ³⁶. A partir de las representaciones SPD, se han propuesto arquitecturas de redes neuronales que mantienen la geometría Riemanniana de los datos de entrada a lo largo de sus capas, permitiendo el aprendizaje compacto de representaciones con la misma geometría inicial ¹⁷. En particular, Huang et al en ¹⁷ presenta un conjunto de tres capas inspiradas en capas de procesamiento de redes de aprendizaje clásicas. Una capa de reducción de dimensión (BiMap), una de regularización (ReEig) y unas últimas para permitir la clasificación (LogEig), ver Figura 2.

1.4. Caracterización oculomotora del Parkinson

Durante la última década, los patrones oculomotores han emergido como un potencial biomarcador temprano que incluso ha demostrado tener correlación con el avance degenerativo producido por la EP ⁵ ⁶ ⁷ ⁹. Sin embargo, cuantificar las alteraciones de los patrones oculomotores relacionadas a la clasificación del Parkinson y su discriminación respectiva con sujetos control que comparten valores demográficos similares, sigue siendo un desafío en esta área de conocimiento. Entre otras, el realizar mediciones en los ojos propias del análisis cinemático, tales como velocidad, amplitud, aceleración y frecuencia, requiere de equipos especializados ⁸. Los métodos de evaluación más precisos del movimiento ocular son mediante electro-oculografías, el sistema de bobina de búsqueda escleral ⁹ y el actualmente más usado, la video-oculografía (VOG). El equipo usado para la medición del VOG, cubre toda la sección ocular, es invasivo, costoso, requiere de calibración y procedimientos de configuración especializados ¹⁰.

En la literatura se han propuesto enfoques con protocolos flexibles de captura, usando secuencias de video ¹⁴ ¹⁶. Por ejemplo, Salazar et al ¹⁴ propuso un método de aprendizaje automático a partir de secuencias de vídeo de una tarea de fijación ocular. Para capturar patrones espacio-

³⁵ Brian Kulis, Mátyás A Sustik e Inderjit S Dhillon. “Low-Rank Kernel Learning with Bregman Matrix Divergences.” En: *Journal of Machine Learning Research* 10.2 (2009).

³⁶ Zhiwu Huang et al. “Projection metric learning on Grassmann manifold with application to video based face recognition”. En: *Proceedings of the IEEE conference on computer vision and pattern recognition*. 2015, págs. 140-149.

temporales, utilizaron cortes 2D volumétricos. A partir de ellos, el método utilizó matrices compactas SPD que resumían representaciones profundas de las alteraciones oculomotoras. Al ser matrices SPD, el método propuesto en este trabajo proyectaba las matrices al plano tangente, para posteriormente clasificarlas utilizando métodos clásicos de aprendizaje de máquina como máquinas de soporte vectorial¹⁴. Sin embargo, realizar tareas de discriminación sobre la variedad de las matrices SPD respetando su geometría no ha sido una tarea fácil, por lo que otros trabajos han propuesto diferentes alternativas para esto. Por ejemplo, combinando múltiples clasificadores en el plano tangente³³ o mapeando las matrices SPD a un espacio vectorial donde se utiliza un método de clasificación basado en representaciones escasas³⁷. Sin embargo, estas estrategias requieren de constantes transformaciones de los datos SPD a espacios de estructura geométrica diferente, pudiendo perder información relevante¹⁷.

A partir de estas premisas, se han propuesto estrategias de aprendizaje profundo sobre el espacio Riemanniano de las matrices SPD. Para esto, Huang et. al. propusieron una red profunda con capas especiales que procesan datos SPD de forma similar a redes de aprendizaje profundo clásicas como las redes densas, mediante técnicas de compactación y regularización, mientras se mantiene la estructura geométrica de la variedad¹⁷. Recientemente, Olmos et al. propusieron una arquitectura híbrida extremo-a-extremo, que parte de representaciones 2D de fijación ocular y propone combinar módulos convolucionales, capas de compactación en matrices SPD y esquemas de aprendizaje profundo Riemanniano para extraer mapas de información, resumir relaciones estadísticas y llevar a cabo un correcto aprendizaje profundo geométrico de matrices SPD¹⁶. Sin embargo, en estos métodos el entrenamiento se mantiene de forma supervisada, i.e, se lleva a cabo con etiquetas definidas. Estas alternativas pueden introducir subjetividad a la predicción del modelo, debido a la subjetividad intrínseca del diagnóstico médico donde se etiquetaron los datos, la cual ha sido cuantificada con un error de más de un 24% en etapas tempranas⁴.

³⁷ Ganggang Dong y Gangyao Kuang. “Target Recognition in SAR Images via Classification on Riemannian Manifolds”. En: *IEEE Geoscience and Remote Sensing Letters* 12.1 (2015), págs. 199-203.

2. PROBLEMA DE INVESTIGACIÓN

El Parkinson es la segunda enfermedad neurodegenerativa con mayor ocurrencia, afectando entre un 2-3 % de la población mundial mayor a 65 años ¹. El Parkinson produce alteraciones motoras, progresivas, pero su detección temprana y seguimiento aún representa un desafío. De hecho, no existe un biomarcador definitivo de la enfermedad. Aunque existen diferentes estudios que proponen metodologías de apoyo al diagnóstico de la enfermedad de Parkinson (EP) ^{38 39 40 41}, estas son comúnmente invasivas o se basan en biomarcadores que se presentan en etapas tardías de la EP. Por ejemplo, diagnósticos basados en el análisis de la marcha o cuantificación del temblor en las manos. Además, en estas etapas los diagnósticos clínicos se reducen a procedimientos observacionales, lo cual provoca que exista una alta subjetividad en el diagnóstico entre diferentes clínicos ⁴.

Diversos estudios han evidenciado una alta relación entre el avance de la enfermedad desde etapas tempranas de la EP con algunos patrones oculomotores ^{20 21 5 6 7}, sin embargo, estos patrones anormales asociados a la EP son difíciles de analizar y cuantificar, al mismo tiempo que costosos de medir. Es así que se requieren diseñar nuevas metodologías que impacten en la

³⁸ Carlotta Caramia, Diego Torricelli, Maurizio Schmid et al. “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.

³⁹ Enas Abdulhay, N Arunkumar, Kumaravelu Narasimhan et al. “Gait and tremor investigation using machine learning techniques for the diagnosis of Parkinson disease”. En: *Future Generation Computer Systems* 83 (2018), págs. 366-373.

⁴⁰ Jerry Ajay, Chen Song, Aosen Wang et al. “A pervasive and sensor-free Deep Learning system for Parkinsonian gait analysis”. En: *2018 IEEE EMBS International Conference on Biomedical Health Informatics (BHI)*. 2018, págs. 108-111.

⁴¹ Kun Hu et al. “Vision-Based Freezing of Gait Detection with Anatomic Patch Based Representation”. En: *Computer Vision – ACCV 2018*. Ed. por C. V. Jawahar et al. Cham: Springer International Publishing, 2019, págs. 564-576.

cuantificación y caracterización ocular para soportar el diagnóstico de la EP y sobre todo que sean posible transferirlos a la rutina clínica. Los métodos computacionales, que involucran el análisis automático de información observacional, han surgido como una alternativa no-intrusiva para soportar el diagnóstico de la enfermedad, así como también para cuantificar patrones atípicos de movimiento a partir de la clasificación de imágenes extraídas de tareas de fijación ocular⁴² 9. Entre ellos, métodos que utilizan matrices SPD han permitido descubrir patrones oculares de la enfermedad en entornos con pocos datos ¹⁴ 16, lo que resulta relevante en ambientes clínicos, donde la disponibilidad de la información es limitada. Estos métodos, sin embargo, son ajustados y optimizados con respecto a reglas supervisadas, que requieren anotaciones por parte de expertos, introduciendo un sesgo inherente en el comportamiento del modelo. Son entonces necesarios métodos que aprovechen la información dada por los descriptores SPD de forma profunda bajo esquemas no supervisados. Lo anterior nos conduce a la siguiente pregunta de investigación:

Pregunta de investigación:

¿Cómo implementar una estrategia no supervisada, teniendo en cuenta la restricción de datos de entrenamiento, para ajustar una arquitectura codificador-decodificador que permita clasificar sujetos control de pacientes con la enfermedad de Parkinson?

⁴² Tomasz Lewandowski y Anna Lewandowska. “Computer-aided processing of the oculomotor signal”. En: *ITM Web of Conferences*. Vol. 21. EDP Sciences. 2018, pág. 00025.

3. OBJETIVOS

Objetivo General

- Desarrollar una arquitectura codificador-decodificador que explote la geometría de matrices compactas SPD para clasificar patrones de fijación ocular en pacientes con Parkinson.

Objetivos Especificos

- Seleccionar un conjunto de datos con patrones oculomotores registrados en secuencias de video y anotados por expertos.
- Construir matrices SPD como descriptores Riemannianos de baja dimensión a partir de mapas de características dados por redes convolucionales pre-entrenadas.
- Implementar una arquitectura codificador-decodificador para la clasificación de patrones oculomotores de fijación por medio de matrices SPD compactas.
- Evaluar el algoritmo desarrollado para la clasificación de patrones oculomotores relacionados con el Parkinson, respecto a métodos del estado del arte basados en aprendizaje supervisado.

4. Conjunto de Datos

Para este trabajo se seleccionó un conjunto de datos de un total de 26 participantes, 13 de ellos diagnosticados con la Enfermedad de Parkinson (EP) y 13 sujetos de control. Los pacientes con EP tenían una edad media de $72,3 \pm 7,4$ años, mientras que los sujetos de control tenían una edad media de $72,2 \pm 6,1$ años. Se utilizó una cámara con una resolución temporal de 60 FPS para capturar la región superior de la cara de los participantes mientras realizaban una tarea de fijación ocular durante 5 segundos. Se recortaron manualmente los ojos individuales a 210×140 píxeles, centrándose en la pupila del primer fotograma para obtener las secuencias de interés. En total, se obtuvieron 5 secuencias de vídeo por participante.

En el conjunto de datos se incluyeron pacientes con diferentes grados de progresión de la EP, categorizados en la escala de valoración Hoehn-Yahr con la ayuda de un fisioterapeuta. Esto con el fin de introducir variabilidad en el problema, incluyendo sujetos con diferente grado de severidad en las anormalidades oculomotoras. De los 13 pacientes Parkinson, cinco pacientes fueron categorizados en la etapa dos, seis pacientes en la etapa tres, y dos pacientes en la etapa cuatro.

Para capturar observaciones del temblor ocular y las anormalidades durante la tarea de fijación, se consideró capturar información espacio-temporal desde un plano bidimensional de observación. Para esto, se codificaron cortes 2d de los videos, como se muestra en la Figura 3. Cada secuencia de vídeo se puede considerar como un volumen $I(x, y, t)^{W, H, N}$ donde la dimensión espacial es $W \times H$ y la temporal N fotogramas. Para obtener los cortes espacio-temporales, se escogieron cuatro direcciones radiales $\theta_i \in \{0^\circ, 45^\circ, 90^\circ, 135^\circ\}$ alrededor del centro del iris y se cortó el volumen de video a través de cada una de estas direcciones (Ver Figura 3). De esta forma, cada corte resultante S_{θ_i} captura información temporal y espacial a través de la dirección θ_i . Como resultado se obtuvo un conjunto de 40 cortes 2D por paciente, correspondientes a los dos ojos, las cinco secuencias de videos, y las 4 direcciones de corte. Estas imágenes 2D que registran sutiles desplazamientos oculares, capturando potenciales alteraciones oculomotoras

relacionadas con la EP.

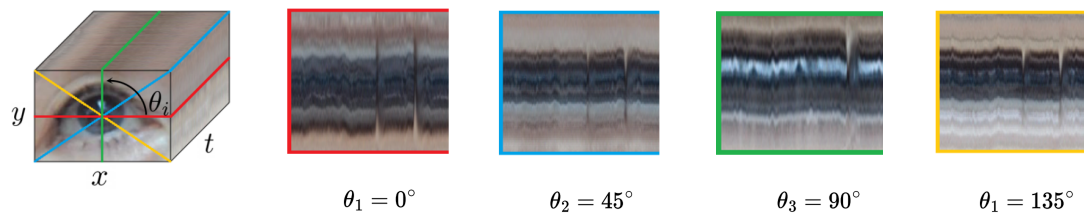


Figura 3. Extracción de cortes temporales en distintos ángulos θ_i partiendo desde vídeos de pacientes ejecutando la tarea de fijación ocular. El objetivo es obtener un posible biomarcador estacionario de la EP.

5. Método Propuesto

Este trabajo propone una estrategia auto-supervisada que logra la cuantificación de patrones oculomotores, desde una representación de matrices SPD compactas, logrando un carácter discriminativo entre pacientes diagnosticados con la EP y pacientes control. Particularmente en el enfoque propuesto utiliza un conjunto de activaciones profundas (*deep features*, de una red pre-entrenada), que proyectan la información de los cortes 2D, que resumen las observaciones de fijación registrados en un video. Este banco de activaciones es codificada en una matriz SPD que guarda las relaciones de segundo orden entre las características. Estas matrices SPD ajustan una arquitectura geométrica autosupervisada, que preserva la información de los datos, mientras aprende relaciones asociadas a la fijación sin la supervisión asociada a las etiquetas. En la Figura 4 se ilustra la arquitectura del método propuesto.

5.1. Descriptor de patrones de fijación

Para cuantificar patrones oculomotores de fijación, los cortes espacio-temporales calculados facilitan la tarea de hallar anomalías de movimiento directamente a partir de secuencias de video. Estos patrones anormales pueden ser observados como frecuencias atípicas que además tienen una alta variabilidad con respecto a cada sujeto registrado. Es por esto que para cada corte espacio-temporal proponemos calcular un descriptor compacto que guarde relaciones entre características profundas de este corte. Para esto, se tomó una red convolucional pre-entrenada en un problema de clasificación de imágenes naturales, la cual brinda características de menor dimensión con suficiente generalidad para descomponer la información de cualquier imagen de entrada, en este caso, de los cortes espacio-temporales. Es así como la red convolucional sirve como extractor de características profundas (activaciones) para cada uno de los cortes proyectados. Los mapas de activación de las primeras capas fueron recuperados, generando un conjunto de características profundas (Deep features). Como resultado se obtuvo un conjunto de

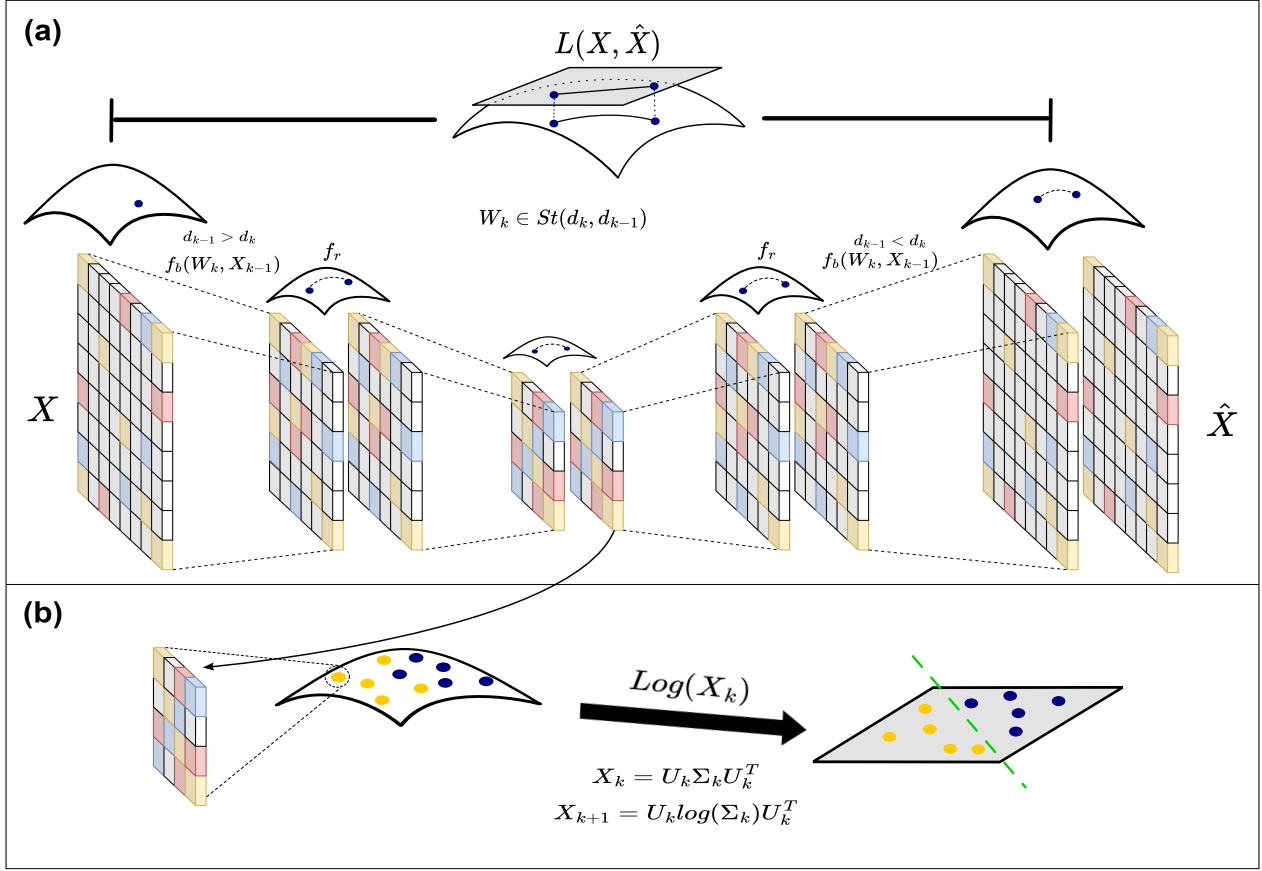


Figura 4. Pipeline del método propuesto. En a) se presenta la arquitectura del autoencoder SPD propuesto. A partir de una matriz SPD de entrada X , el codificador (*encoder*) proyecta la matriz a un embebido X_K tras aplicar K capas de aprendizaje Riemanniano profundo. El decodificador reconstruye este embebido en una matriz SPD \hat{X} . El aprendizaje del modelos se rige por la perdida de reconstrucción $L(X, \hat{X})$. b) Posterior al entrenamiento, se proyectan los embebidos SPD al plano tangente (espacio Euclidean) para ser clasificados utilizando las etiquetas dadas por un experto.

C activaciones (codificadas como imágenes 2D) $\{D_1, D_2, \dots, D_C\}$, donde cada imagen D_i tiene una dimensión espacial de $W \times H$. Posteriormente para resumir la información codificada en estas características se calculó una matriz de covarianza. A partir del conjunto de características $\{D_1, D_2, \dots, D_C\}$ esta matriz es dada por $M = [\gamma_{ij}]_{i=1, j=1}^{C, C}$, donde $\gamma_{ij} = E[(D_i - E[D_i])(D_j - E[D_j])]$ representa la covarianza entre la característica D_i y D_j . Como resultado se obtiene una matriz compacta de dimensión $C \times C$, mucho menor que el banco de características de entrada (dimensión $C \times W \times H$). El proceso para el calculo de este descriptor se ilustra en la Figura 5.

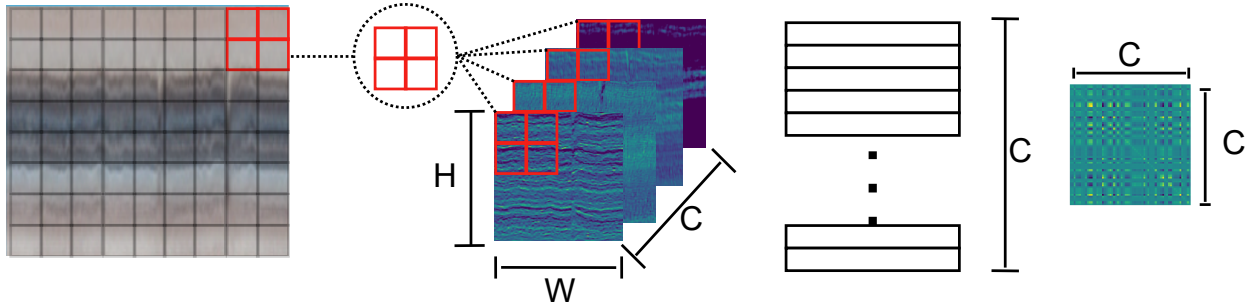


Figura 5. Calculo de descriptor compacto SPD. Cada corte fue descrito por mapas de activación extraídos de una red convolucionales pre-entrenada y posteriormente fue compactado en una matriz de covarianza.

5.2. Autoencoder SPD

En este trabajo buscamos proponer una arquitectura autosupervisada que reciba una matriz SPD como entrada y reconstruya una matriz SPD, preservando la geometría Riemanniana donde estos datos coexisten. A continuación se describe la construcción del codificador (5.2.1), decodificador (5.2.2), y la respectiva estrategia de aprendizaje Riemanniano (5.2.3).

5.2.1. Codificador de matrices SPD Para codificar matrices SPD son necesarias capas de aprendizaje profundo Riemanniano que reduzcan la dimensión de entrada mientras se alcanzan representaciones más discriminativas con respecto al problema. En la literatura existen redes de aprendizaje profundo sobre matrices SPD que han tratado con esto. En este trabajo utilizaremos la estructura de red SPD propuesta por *Huang et.al.* en ¹⁷. Esta red propone dos capas de aprendizaje, la capa *BiMap* y *ReEig*, que permiten codificar representaciones SPD en nuevas representaciones aún más compactas teniendo en cuenta las restricciones geométricas. A continuación se describen a detalle estas capas:

- **BiMap:** El comportamiento de esta capa es similar a la capa densa de las redes neuronales clásicas ya que está diseñada para generar representaciones de menor dimensión. En este caso, a partir de una matriz SPD X_{k-1} , la capa *BiMap* genera una nueva matriz SPD de menor dimensionalidad X_k , por medio de un mapeo bilineal f_b . Este mapeo se describe

como:

$$X_k = f_b^{(k)}(X_{k-1}, W_k) = W_k X_{k-1} W_k^T \quad (3)$$

Donde f_b recibe como entrada $X_{k-1} \in \text{Sym}_{d_{k-1}}^+$, es decir, una matriz SPD de dimensión d_{k-1} . Esta función depende también de una matriz de peso $W_k \in \mathbb{R}^{d_k \times d_{k-1}}$, con la cual se calcula $X_k \in \text{Sym}_{d_k}^{++}$, la nueva matriz SPD de dimensión d_k . Al igual que una red clásica, la matriz de peso o matriz de transformación W_k es entrenable. Adicionalmente, W_k se condiciona a ser una matriz de filas linealmente independientes (*row full-rank*) para asegurar que X_k sea SPD y así poder mantener la geometría de los descriptores durante el entrenamiento de la red ⁴³. Sin embargo, la optimización sobre estos pesos lleva a problemas de factibilidad. Para esto, se ha propuesto considerar los pesos W_k como elementos de una variedad de Stiefel no compacta $St(d_k, d_{k-1}) = \{W_k \in \mathbb{R}^{d_k \times d_{k-1}} | W_k W_k^T = I_{d_k}\}$ ⁴⁴.

- **ReEig**: Inspiradas en la ReLu, estas capas introducen no-linealidad para ayudar al aprendizaje de la red. Esto a partir del calculo

$$X_k = f_r^{(k)}(X_{k-1}) = U_{k-1} \max(\epsilon I, \Sigma_{k-1}) U_{k-1}^T \quad (4)$$

Donde, U_{k-1} y Σ_{k-1} son la respectiva descomposicion espectral (eigendescomposición) de la matriz de entrada $X_{k-1} = U_{k-1} \Sigma_{k-1} U_{k-1}^T$, siendo U_{k-1} los eigenvectores, Σ_{k-1} una matriz diagonal con los respectivos eigenvalores, ϵ es un hiperparámetro de rectificación e

⁴³ Zhiwu Huang y Luc Van Gool. “A riemannian network for spd matrix learning”. En: *Proceedings of the AAAI conference on artificial intelligence*. Vol. 31. 1. 2017.

⁴⁴ Vincent Arsigny et al. “Geometric means in a novel vector space structure on symmetric positive-definite matrices”. En: *SIAM journal on matrix analysis and applications* 29.1 (2007), págs. 328-347.

I es la matriz identidad de dimensión d_{k-1} . Aquí el cálculo de $\max(\epsilon I, \Sigma_{k-1})$ es dado por

$$A(i, i) = \begin{cases} \Sigma_{k-1}(i, i) & \text{si } \Sigma_{k-1}(i, i) > \epsilon \\ \epsilon & \text{si } \Sigma_{k-1}(i, i) \leq \epsilon \end{cases} \quad (5)$$

De esta manera, se analiza cada autovalor de la matriz SPD X_{k-1} y se iguala a ϵ si $\Sigma_{k-1}(i, i) < \epsilon$. Esto, además de agregar no linealidad al modelo, en busca de la conservación de la geometría de las matrices, asegura que se mantengan autovalores positivos.

Con estas dos capas, *BiMap*, que genera representaciones mas compactas y *ReEig*, que mantiene las matrices positivas, se define el bloque *BiRe*, que se compone de una capa *BiMap* seguida de una *ReEig* $f_r(WXW^T)$. Similar a las redes clásicas donde los bloques de procesamiento se componen de capas convolucionales seguidas de capas de rectificación, los bloques *BiRe* permiten codificar y compactar los datos SPD de entrada. El ajuste de los pesos W_k permite la codificación de una matriz SPD mas compacta.

5.2.2. Decodificador de matrices SPD Una de las principales contribuciones de este trabajo se encuentra en la construcción de bloques decodificadores que permitan la reconstrucción de matrices SPD compactas, y que junto con el proceso de codificación permitan un aprendizaje auto-supervisado. Esta sección introduce las estructuras de la red propuesta que permiten procesar el embebido SPD generado por el codificador de forma que se aumente progresivamente su dimensión hasta generar una matriz SPD de la misma dimensión que la matriz de entrada de la red. Para esto, se busca una capa que realice la tarea inversa de la capa *BiMap*. En el trabajo de Huang et al. ¹⁷, donde se introducen los componentes *BiMap*, se establece que estos no pueden ser utilizados en tecnicas de sobremuestreo (upsampling) planteando que es necesario que $W_k \in \mathbb{R}^{d_k \times d_{k-1}}$ y $d_k < d_{k-1}$. Puesto que para la decodificación de matrices SPD se tiene lo contrario, $d_k > d_{k-1}$, a continuación demostramos que bajo esta condición, el cálculo de f_b se mantiene factible. Permitiendo generar matrices simétricas semi-definidas positivas de mayor

dimensión.

Proposición 1: Dada una matriz simétrica definida positiva X_{k-1} de dimensión d_{k-1} ($\in Sym_{d_{k-1}}^{++}$) y una matriz $W_k \in St(d_k, d_{k-1})$, donde $d_k < d_{k-1}$, se tiene que $f_b^{(k)}(X_{k-1}, W_k)$ es una matriz simétrica semi-definida positiva de dimensión d_k ($\in Sym_{d_k}^{++}$).

Demostración.

Sea $X_{k-1} \in Sym_{d_{k-1}}^{++}$ y sea X_k la matriz resultante al aplicar $f_b^{(k)}(X_{k-1}, W_k) = W_k X_{k-1} W_k^T$, es decir, $X_k = W_k X_{k-1} W_k^T$. Veamos primero que X_k es simétrica. Esto es cierto pues $X_k = X_k^T$ ya que

$$\begin{aligned} X_k^T &= (W_k X_{k-1} W_k^T)^T \\ &= (W_k^T)^T X_{k-1}^T W_k^T \\ &= W_k X_{k-1} W_k^T \\ &= X_k, \end{aligned}$$

aquí se utilizó la propiedad matricial $(A^T)^T = A$ y dado que X_{k-1} es simétrica, $X_{k-1} = X_{k-1}^T$. Ahora demostraremos que X_k es semi-definida. Para esto, demostraremos que los autovalores de X_k son no nulos. Dado que $X_{k-1} \in Sym_{d_{k-1}}^{++}$ admite una descomposición espectral $X_{k-1} = U_{k-1} \Sigma_{k-1} U_{k-1}^T$, donde U_{k-1} y Σ_{k-1} representan la matriz de autovectores y autovalores respectivamente. Usando esto, se tiene que

$$\begin{aligned} X_k &= W_k X_{k-1} W_k^T \\ &= W_k U_{k-1} \Sigma_{k-1} U_{k-1}^T W_k^T. \end{aligned}$$

Dado un vector no nulo $v \in \mathbb{R}^{d_k}$, se tiene que

$$\begin{aligned}
v^T X_k v &= v^T W_k X_{k-1} W_k^T v \\
&= v^T W_k U_{k-1} \Sigma_{k-1} U_{k-1}^T W_k^T v \\
&= Z \Sigma_{k-1} Z^T,
\end{aligned}$$

siendo $Z = v^T W_k U_{k-1} \in \mathbb{R}^{d_{k-1}}$, se tiene que $Z \Sigma_{k-1} Z^T = \sum_{i=1}^{d_{k-1}} Z_i^2 \lambda_i$, siendo $Z = [Z_i]_{i=1}^{d_{k-1}}$ y λ_i los valores de la diagonal Σ_{k-1} . Dado que $\lambda \geq 0$ para todo $i \in \{1, \dots, d_{k-1}\}$ por ser X_{k-1} positiva, se tiene que $v^T X_k v = Z \Sigma_{k-1} Z^T \geq 0$. Por tanto, X_k es semi-definida positiva.

Esta proposición plantea la oportunidad de utilizar la capa f_b no solo como un método de agrupamiento para matrices SPD, sino también como una técnica de upsampling. Esta capa cuenta con pesos W entrenables y, en combinación con la rectificación de la capa f_r , garantiza que la matriz resultante sea positiva definida.

5.2.3. Esquema de aprendizaje Para aprender los parámetros de las capas BiMap del codificador y decodificador, en este trabajo se estudia el uso de dos funciones de pérdida que permiten el ajuste de las representaciones embebidas a partir de una tarea de reconstrucción. En primer lugar, se utilizó el error medio cuadrático (MSE), para el cálculo de la pérdida

$$l_{MSE}(X, \hat{X}) = \frac{1}{n^2} \|X - \hat{X}\|_2^2. \quad (6)$$

Esta pérdida es ampliamente utilizada en modelos autosupervisados, especialmente en autoencoders ⁴⁵. Con el error medio cuadrático se busca estimar la diferencia promedio existente entre la variable real y la generada. Para nuestro problema, esta estimación se traduce en comparar las matrices X y \hat{X} punto a punto, obteniendo una pérdida de reconstrucción enfocada en la generación de una matriz numéricamente similar a la de entrada. Teniendo en cuenta las con-

⁴⁵ Shuangfei Zhai y Zhongfei Zhang. "Semisupervised autoencoder for sentiment analysis". En: *Proceedings of the AAAI Conference on Artificial Intelligence*. Vol. 30. 1. 2016.

diciones geométricas de nuestro enfoque, se propuso una función de pérdida que considera el espacio curvo SPD donde se encuentran los datos de entrada (X) y los reconstruidos (\hat{X}). Para esto, se utilizó la función Riemanniana logaritmo $f_l : Sym_{++}^n \rightarrow T_{I_n} Sym_{++}^n$, la cual proyecta datos SPD al plano tangente en la matriz identidad I_n . Este espacio tangente se puede ver como un espacio Euclideo (plano), por lo que el cálculo de esta proyección permite aplicar operaciones Euclidianas a los datos SPD. Específicamente, la función logaritmo está dada por

$$X_k = f_l^{(k)}(X_{k-1}) = \text{Log}(X_{k-1}) = U_{k-1} \text{Log}(\Sigma_{k-1}) U_{k-1}^T, \quad (7)$$

donde, por eigendescomposición $X_{k-1} = U_{k-1} \Sigma_{k-1} U_{k-1}^T$. Mientras que $\text{Log}(\Sigma_{k-1})$ aplica la operación logaritmo a cada uno de los puntos de la diagonal de Σ_{k-1} . La función de pérdida planteada en 8 considera el hecho de que la matriz de entrada X puede tener eigenvalores muy cercanos a 0, aumentando el riesgo de tener que lidiar con errores numéricos, por ello se propone utilizar el error cuadrático medio de dos matrices SPD regularizadas por la capa *ReEig* y proyectadas en el plano tangente por la capa *LogEig*:

$$L_{\text{Log-MSE}} = \frac{1}{n^2} \|f_l(f_r(X)) - f_l(\hat{X})\|_2^2 \quad (8)$$

Donde f_r y f_l son las capas *ReEig* y *LogEig* respectivamente.

Una vez establecida la pérdida, a partir de l_{MSE} o $L_{\text{Log-MSE}}$, el proceso de retropropagación del error para el aprendizaje se sigue de la misma manera, como se describirá a continuación.

Partiendo de una pérdida cualquiera L , el algoritmo de retropropagación utiliza la regla de la cadena, computando los gradientes como:

$$\frac{\partial L^{(k)}}{\partial \Theta_k} = \frac{\partial L^{(k+1)}}{\partial X_k} \frac{\partial f^{(k)}}{\partial \Theta_k}, \quad (9)$$

donde Θ_k , $L^{(k)}$, y $f^{(k)}$ representan los parámetros, la pérdida y la función aplicada en la k -ésima capa, respectivamente. A partir de esto, el procedimiento convencional para actualizar

iterativamente los parámetros Θ_k , con el objetivo de alcanzar un mínimo local, sigue el algoritmo del gradiente descendiente:

$$\Theta_k = \Theta_k - \alpha \frac{\partial L^{(k)}}{\partial \Theta_k}, \quad (10)$$

donde α es un paso que controla el descenso del gradiente. Sin embargo, surgen dos problemas al utilizar estos gradientes euclidianos y retropropagación tradicional dentro de las capas Riemannianas. El primero es que la egiendescomposición de las capas *ReEig* y *LogEig* no ha sido bien resuelta para la retropropagación tradicional. Y el segundo consiste en que la actualización de los pesos de conexión en las capas *BiMap* no produce matrices ortonormales válidas ¹⁷.

Para enfrentar el primer problema, es necesario computar los gradientes en las capas Eig (*ReEig* y *LogEig*), teniendo en cuenta que nos interesa computar $\frac{\partial L^{(k)}}{\partial X_{k-1}}$. Para esto, se introduce una capa ficticia k' luego de cada capa Eig. Resultando un gradiente dado por:

$$\frac{\partial L^{(k)}}{\partial X_{k-1}} = 2U_{k-1} \left(P^T \circ \left(U_{k-1}^T \frac{\partial L^{(k')}}{\partial U_{k-1}} \right)_{sym} \right) U_{k-1}^T + U_{k-1} \left(\frac{\partial L^{(k')}}{\partial \Sigma_{k-1}} \right)_{diag} U_{k-1}^T,$$

donde \circ es el producto de hadamard, $A_{sym} = \frac{1}{2} (A + A^T)$, A_{diag} en una matriz diagonal con los elementos en la diagonal de A y

$$P(i, j) = \begin{cases} \frac{1}{\lambda_i - \lambda_j}, & i \neq j. \\ 0, & i = j \end{cases},$$

donde λ_i son los eigenvalores Σ_{k-1} . Ahora, los gradientes de la capa ficticia $\frac{\partial L^{(k')}}{\partial U_{k-1}}$ y $\frac{\partial L^{(k')}}{\partial \Sigma_{k-1}}$ dependen de la respectiva capa Eig.

- En el caso de *ReEig*, el gradiente es:

$$\frac{\partial L^{(k')}}{\partial U_{k-1}} = 2 \left(\frac{\partial L^{(k+1)}}{\partial X_k} \right)_{sym} U_{k-1} \text{ máx}(\epsilon I, \Sigma_{k-1}),$$

$$\frac{\partial L^{(k')}}{\partial \Sigma_{k-1}} = QU_{k-1}^T \left(\frac{\partial L^{(k+1)}}{\partial X_k} \right)_{sym} U_{k-1},$$

donde Q es la matriz diagonal

$$Q(i, i) = \begin{cases} 1, & \Sigma_{k-1}(i, i) > \epsilon, \\ 0, & \Sigma_{k-1}(i, i) \leq \epsilon. \end{cases}.$$

- En el caso de *LogEig*, el gradiente es:

$$\begin{aligned} \frac{\partial L^{(k')}}{\partial U_{k-1}} &= 2 \left(\frac{\partial L^{(k+1)}}{\partial X_k} \right)_{sym} U_{k-1} \log(\Sigma_{k-1}), \\ \frac{\partial L^{(k')}}{\partial \Sigma_{k-1}} &= \Sigma_{k-1}^{-1} U_{k-1}^T \left(\frac{\partial L^{(k+1)}}{\partial X_k} \right)_{sym} U_{k-1}. \end{aligned}$$

Para el segundo problema, el aprendizaje depende del cálculo de la dirección de descenso más marcada dentro de la variedad de *Stiefel* donde pertenecen los pesos de conexión *BiMap*¹⁷. Dado un peso de conexión $W_k \in St(d_k, d_{k-1})$ en el t -ésimo paso del algoritmo, su gradiente euclidiano con respecto a la función de pérdida $\nabla L^{(k)} W_k^t$ (ver ecuación 9) no es consistente con la geometría de la variedad de *Stiefel*. Por tanto, es necesario calcular la componente tangente a $St(d_k, dk - 1)$, definiéndose como la resta del gradiente euclidiano $\nabla L^{(k)} W_k^t$ y la componente normal a la variedad: $\tilde{\nabla} L^{(k)} W_k^t = \nabla L^{(k)} W_k^t - \nabla L^{(k)} W_k^t (W_k^t)^\top W_k^t$. Ahora, la componente tangente $\tilde{\nabla} L^{(k)} W_k^t$ se puede ver como la dirección para actualizar el peso de conexión W_k^t . Finalmente, se utiliza la operación de retracción Γ , definida como la matriz Q de la descomposición QR ⁴⁶, para mapear este peso en $St(d_k, dk - 1)$. Obteniendo un peso actualizado W_k^{t+1} utilizando una tasa de aprendizaje α y una variación del algoritmo del descenso de gradiente definido en 10, mediante la siguiente ecuación:

⁴⁶ P-A Absil, Robert Mahony y Rodolphe Sepulchre. *Optimization algorithms on matrix manifolds*. Princeton University Press, 2009.

$$W_k^{t+1} = \Gamma \left(W_k^t - \alpha \tilde{\nabla} L^{(k)} W_k^t \right), \quad (11)$$

siendo α el paso del descenso del gradiente en esta variedad. Como resultado se tiene un algoritmo de retropropagación que permite actualizar los parámetros de la red propuesta respetando las condiciones geométricas.

5.3. Clasificación de embebidos

Luego de realizar aprendizaje no supervisado para ajustar las representaciones embebidas del autoencoder y con el fin de evaluar la capacidad discriminatoria de las representaciones compactas, se realizó un análisis del espacio embebido resultante de nuestro modelo. Para esto se tomaron los resultados de la ultima capa del codificador.

A partir de estos embebidos, propusimos diferentes algoritmos de clasificación supervisados para analizar la capacidad del modelo de generar representaciones discriminativas en el problema de clasificación de la enfermedad de Parkinson, utilizando estos embebidos y las etiquetas dadas por los expertos médicos. Debido a que estas representaciones embebidas están en la variedad geométrica, se realizó previamente una proyección al plano tangente en la identidad utilizando el operador logaritmo y posteriormente las matrices resultantes fueron vectorizadas, filtrando por los datos de la matriz triangular superior para evitar redundancia, ver figura 4 b). En este caso, se proyectan los embebidos del autoencoder ajustado utilizando la funcion $Log(X_k)$ donde X_k es el embebido. Esta funcion las proyecta en el plano tangente donde serán clasificados por los siguientes algoritmos:

- **Bosques aleatorios:** Consiste en una combinación aleatoria de clasificadores *arboles de decisión*. Particularmente cada *árbol de decisión* representa ramas aprendidas con umbrales de partición τ_m a lo largo cada valor en la matriz embebida SPD X_k (triangular superior). Los árboles que representan de forma independiente la hipótesis de la partición de las la matriz embebidoa SPD X se construyeron a partir de un análisis óptimo y dis-

criminante usando la estrategia *CART - Classification and Regression Trees*. Entonces, se obtiene una clasificación para cada árbol en el bosque aleatorio, dando como etiqueta predicha, la etiqueta de mayor recurrencia entre los árboles aprendidos ⁴⁷.

- **Maquina de soporte vectorial:** El algoritmo es utilizado para encontrar un hiperplano mediante la maximización del margen entre una clase y otra. En algunos casos donde el conjunto de datos no es linealmente separables, se emplea una función de "kernel" que transforma los datos de entrada en un nuevo conjunto (usualmente de mayor dimensión), donde se logra la separación de los datos ⁴⁸. En este trabajo, las matrices embebidas $\log(X_k)$ son proyectadas a una máquina de soporte vectorial con un kernel rbf (*Radial basis function*), popular por su comportamiento similar al vecino mas cercano.
- **Regresión logística:** es un modelo de regresión que se ajusta a un problema de clasificación binarios a partir de estimar las probabilidades para cada evento (dato) de entrada. En esta regresión, se utiliza una cross entropía binaria de las etiquetas reales en función de las probabilidades predichas. La ecuación de regresión es $\ln(\frac{P}{1-P}) = \beta_0 + \beta_1 X$, donde P representa la probabilidad de ocurrencia del evento, β_0 es la intercepción y β_1 son los coeficientes de regresión a ajustar ⁴⁹. Esta regresión logística es de principal interés para nuestro trabajo y que permite medir la capacidad del descriptor aprendido sin etiquetas para separar clases usando planos lineales, siendo uno de los componentes mas simples de separación. Conservando un mayor esfuerzo en la creación de las matrices SPD, y no en el entrenamiento de los algoritmos clasificadores para no tener una fuerte dependencia de las etiquetas.

⁴⁷ Mahesh Pal. "Random forest classifier for remote sensing classification". En: *International journal of remote sensing* 26.1 (2005), págs. 217-222.

⁴⁸ Lipo Wang. *Support vector machines: theory and applications*. Vol. 177. Springer Science & Business Media, 2005.

⁴⁹ Michael P LaValley. "Logistic regression". En: *Circulation* 117.18 (2008), págs. 2395-2399.

5.4. Configuración experimental

Para la evaluación del enfoque propuesto se consideraron diferentes conjuntos de datos de matrices SPD construidos a partir de mapas de activación extraídos de redes convolucionales pre-entrenadas. A partir del modelo propuesto se consideraron diferentes arquitecturas, variando la profundidad de la red y función de pérdida con el fin de analizar el impacto de estas componentes en las representaciones alcanzadas. Por último el modelo fue evaluado sobre una población de pacientes Parkinson y sujetos control.

5.4.1. Mapas de activación Dos diferentes arquitecturas convolucionales pre-entrenadas fueron usadas de forma independiente para la extracción de los mapas de activación y compactación en descriptores SPD (ver figura 5). Cada una de las arquitecturas escogidas fue previamente entrenada en un problema general de clasificación de imágenes naturales ⁵⁰. A continuación se describe de forma general estas arquitecturas:

- **VGG-19** ⁵¹ es una arquitectura clásica con un total de 19 capas. Para la extracción de mapas de activación se consideraron representaciones de bajo nivel provenientes de la salida del primer bloque de agrupamiento. En total se obtuvieron 64 mapas de activación, cada uno de ellos de dimensión 112×112 . La arquitectura VGG tiene como principal ventaja la implementación original de convolución con Kernels enfocados en características estrictamente locales (dimension de kernel de 3×3). En este sentido, el banco de características logra una descomposición de los diferentes componentes espectrales de los cortes de ojo proyectados.

⁵⁰ Alex Krizhevsky, Ilya Sutskever y Geoffrey E Hinton. “Imagenet classification with deep convolutional neural networks”. En: *Communications of the ACM* 60.6 (2017), págs. 84-90.

⁵¹ Karen Simonyan y Andrew Zisserman. “Very deep convolutional networks for large-scale image recognition”. En: *arXiv preprint arXiv:1409.1556* (2014).

- **Xception**⁵² es una arquitectura de red neuronal convolucional profunda que involucra convoluciones separables en profundidad. Se obtuvieron 64 mapas de activación de dimensión 147×147 obtenidos del primer bloque de agrupamiento. Esta arquitectura es reciente e involucra mecanismos para la reducción computacional, como los cuello de botella (con Kernels de tamaño 1×1), logrando ser mas profunda sin perder capacidades de ajuste.

5.4.2. Arquitectura codificador-decodificador Para clasificar cada uno de los conjuntos de matrices SPD descritos anteriormente se proponen diferentes arquitecturas. En primer lugar se consideró variar la profundidad del autoencoder SPD (ASPD). Para esto, se consideró diferente cantidad de bloques *BiRe* en el codificador y usando de forma simétrica la misma cantidad de capas en el decodificador. En resumen seCon un total de 3 configuraciones planteadas:

- **ASPD-1.** Se compone de un bloque *BiRe* codificador que reduce la matriz de entrada $X \in \mathbb{R}^{64 \times 64}$ a $X_1 \in \mathbb{R}^{32 \times 32}$. Posteriormente se utiliza un bloque *BiRe* en el decodificador para generar $\hat{X} \in \mathbb{R}^{64 \times 64}$ a partir de X_1 .
- **ASPD-2.** Se compone de dos bloques *BiRe* en el codificador que reducen la matriz de entrada $X \in \mathbb{R}^{64 \times 64}$ a $X \rightarrow X_1 \in \mathbb{R}^{32 \times 32} \rightarrow X_2 \in \mathbb{R}^{16 \times 16}$. Posteriormente se utilizan dos bloques *BiRe* en el decodificador para generar $X_2 \rightarrow X_3 \in \mathbb{R}^{32 \times 32} \rightarrow \hat{X} \in \mathbb{R}^{64 \times 64}$.
- **ASPD-3.** Se compone de tres bloques *BiRe* en el codificador que reducen la matriz de entrada $X \in \mathbb{R}^{64 \times 64}$ a $X \rightarrow X_1 \in \mathbb{R}^{32 \times 32} \rightarrow X_2 \in \mathbb{R}^{16 \times 16} \rightarrow X_3 \in \mathbb{R}^{8 \times 8}$. Posteriormente se utilizan tres bloques *BiRe* en el decodificador para generar $X_3 \rightarrow X_4 \in \mathbb{R}^{16 \times 16} \rightarrow X_5 \in \mathbb{R}^{32 \times 32} \rightarrow \hat{X} \in \mathbb{R}^{64 \times 64}$.

Estas arquitecturas fueron ajustadas con una tasa de aprendizaje de $\alpha = 0,001$ y un tamaño de batch de 32. El entrenamiento de cada modelo se realizó en 50 épocas. Con respecto a la

⁵² François Chollet. “Xception: Deep learning with depthwise separable convolutions”. En: *Proceedings of the IEEE conference on computer vision and pattern recognition*. 2017, págs. 1251-1258.

función de pérdida, se evaluó el impacto de ajustar los modelos utilizando la pérdida L_{MSE} o $L_{MSE-Log}$, las cuales se describieron en las ecuaciones 4 y 7.

5.4.3. Validación El enfoque propuesto fue validado usando un esquema de validación cruzada 5k-fold. Cada fold contenía 21 pacientes de entrenamiento ($21 \times 40 = 840$ muestras) y 5 pacientes de test ($5 \times 40 = 200$ muestras), manteniendo como mínimo 2 pacientes con EP y 2 pacientes control en el conjunto de test. Hay un solo fold que posee 20 pacientes de entrenamiento y 6 de test, conteniendo clases balanceadas en ambos conjuntos. Los modelos de autoencoder SPD propuestos fueron ajustados de manera independiente con el conjunto de datos de cada fold.

5.4.4. Clasificación de embebidos Una vez entrenados los modelos de autoencoders, se extrajeron los embebidos de la capa 1, 2 y 3 para el modelo ASPD-1, ASPD-2 y ASPD-3 respectivamente. Posteriormente, para medir la capacidad de discriminación, se utilizaron los algoritmos de bosques aleatorios (Random Forest, RF); máquina de soporte vectorial (Support vector machine, SVM); y regresión logística (Logistic Regression, LG) para la clasificación de los embebidos. Estos modelos fueron ajustados para cada fold y evaluados sobre el conjunto de test del fold correspondiente. Para medir el rendimiento se calculó el promedio de exactitud (Accuracy, denotado como Acc), sensibilidad (Sensitivity, denotado como Sen) y puntaje F1 (F1-score, denotado como $F1$), para cada fold en los respectivos clasificadores.

6. EVALUACIÓN Y RESULTADOS

La estrategia de autoencoder SPD propuesta fue evaluada en la tarea de reconstrucción de patrones de fijación ocular asociados a la enfermedad de Parkinson. Esta tarea de pretexto pretende involucrar componentes autosupervisados con respecto a datos de entrada, sin tener en cuenta sesgos que puedan estar asociados a las etiquetas. Una vez ajustada la arquitectura para reconstruir matrices SPD, un paso siguiente consiste en utilizar las matrices embebidas como descriptores para discriminar entre observaciones de una población con Parkinson y una población control. El primer enfoque para la validación fue validar la capacidad, de diferentes versiones de la arquitectura, de reconstruir matrices. Como métrica de reconstrucción se utilizó el clásico MSE que permite controlar los valores atípicos en la reconstrucción, así como también el MSE-Log que hace una proyección de las matrices SPD en su plano Euclidiano. Se diseñaron dos conjuntos de datos SPD a partir de mapas de activación extraídos de arquitecturas convolucionales preentrenadas; la VGG19 y la Xception. A partir de esto, se diseñaron tres modelos ASPD-1BiRe, ASPD-2BiRe y ASPD-3BiRe, variando el número de n bloques BiRe en el codificador y n bloques BiRe en el decodificador. Esto con el fin de analizar el impacto de arquitecturas más profundas en el desempeño método propuesto y el aprendizaje de representaciones más discriminativas. Para el ajuste de dichos modelos, en primera instancia se consideró el ajuste de las representaciones usando una pérdida clásica de reconstrucción L_{MSE} , alcanzando los errores de reconstrucción mostrados en la tabla 1.

En la tabla 1 se observa que la mejor configuración para minimizar la función de pérdida L_{MSE} corresponde al modelo ASPD-1BiRe que utiliza como entrada los datos provenientes de la VGG19. También se observa como la calidad de la reconstrucción es inversamente proporcional a la profundidad. Sin embargo, al hallar el MSE de las matrices regularizadas y proyectadas al plano tangente (MSE-log), se obtiene que la configuración ASPD-1BiRe usando como entrada datos provenientes de la Xception minimiza más la pérdida de reconstrucción.

En un segundo experimento de reconstrucción, las arquitecturas consideradas fueron evaluadas

Data	Model	L_{MSE}	MSE-Log
VGG19	ASPD-1BiRe	0.505	0.47
	ASPD-2BiRe	0.52	0.594
	ASPD-3BiRe	0.541	0.658
Xception	ASPD-1BiRe	2.674	0.316
	ASPD-2BiRe	2.652	0.318
	ASPD-3BiRe	2.851	0.325

Tabla 1. Resultados errores de reconstrucción L_{MSE} . Errores de reconstrucción de los modelos propuestos variando el número de bloques BiRe. Evaluados sobre los dos conjunto de datos SPD (Data) construidos utilizando como perdida el error medio cuadrático (MSE).

con una función de perdida que hace una consideración sobre la geometría de los datos, *i.e.*, que estos no se encuentran mapeados en un espacio Euclidiano. Para ello, las arquitecturas fueron ajustadas con respecto a una función de perdida Riemanniana $L_{MSE-Log}$. Los resultados correspondientes a las variaciones del conjunto de datos y profundidad de la arquitectura ajustada con función de perdida Riemanniana se encuentran descritos en la tabla 2

Data	Model	MSE	$L_{MSE-log}$
VGG19	ASPD-1BiRe	0.739	0.463
	ASPD-2BiRe	0.625	0.562
	ASPD-3BiRe	0.618	0.628
Xception	ASPD-1BiRe	3.914	0.27
	ASPD-2BiRe	3.884	0.287
	ASPD-3BiRe	3.936	0.304

Tabla 2. Resultados errores de reconstrucción $L_{MSE-log}$. Errores de reconstrucción de los modelos propuestos variando el número de bloques BiRe. Evaluados sobre los dos conjunto de datos SPD (Data) construidos utilizando como perdida el error medio cuadrático de una proyección en el plano tangente (MSE-Log).

En la tabla 2 se observa que la mejor configuración para minimizar la función de perdida $L_{MSE-log}$ corresponde al modelo ASPD-1BiRe que utiliza como entrada los datos provenientes de la Xception. Esta tabla también muestra que la profundidad de la arquitectura tiende a empeorar la métrica de reconstrucción dada por la función de perdida. Además, en conjunto con la tabla 1, se observa que el MSE es menor para todas las configuraciones que usan el conjunto de datos dado por la VGG19. Una exploración un poco más detallada de los datos mostró que

esta diferencia en la métrica de reconstrucción puede estar dada por la cantidad de autovalores que se pierden en la regularización de las capas ReEig, ya que aproximadamente el 45 % de los autovalores del conjunto de datos proveniente de la Xception, están siendo regularizados en la primera capa ReEig.

En un segundo enfoque experimental, se validó la capacidad de las representaciones autoajustadas para codificar patrones de fijación con capacidad de discriminación entre la población de Parkinson y la de control. En este caso, desde las arquitecturas entrenadas se obtuvieron las representaciones embebidas correspondientes al conjunto de datos de entrenamiento. Posteriormente, los algoritmos de clasificación como Bosques aleatorios, Máquina de soporte vectorial y Regresión logística fueron seleccionados para clasificar estos embebidos. Los resultados obtenidos en la clasificación de embebidos obtenidos en el entrenamiento sobre el conjunto VGG19 y Xception se muestran en la Tabla 3 y Tabla 4 respectivamente.

En la Tabla 3, para los modelos entrenados el conjunto de datos VGG19 y con L_{MSE} , se observa que la mejor configuración con un 0.923 de exactitud fue utilizando una regresión logística (LR) y la arquitectura ASPD-1BiRe. Mientras que para los métodos que utilizaron $L_{MSE-log}$, se observa que el mejor método logró una exactitud de 0.903, un 0.02 menos que el mejor método usando L_{MSE} . Este resultado se obtuvo utilizando el mismo conjunto de datos (VGG19), arquitectura (ASPD-1BiRe) y clasificador (LR) que el mejor método ajustado con la función de pérdida L_{MSE} .

En la Tabla 4, para los modelos entrenados con el conjunto de datos Xception y con L_{MSE} , se observa que la mejor configuración con un 0.883 de exactitud fue un clasificador de bosques aleatorios (RF) con una arquitectura ASPD-3BiRe. Por otra parte, para los modelos ajustados con la función de pérdida $L_{MSE-log}$, se observa que el mejor método logró una exactitud de 0.86, manteniendo 0.02 menos que el mejor método usando L_{MSE} . Este fue el resultado de usar la arquitectura ASPD-2BiRe y una regresión logística (LR) de clasificador.

Ambas tablas muestran que la regresión logística supera los otros clasificadores en la mayoría de los casos, lo que implica que se mantiene una linealidad intrínseca en el espacio embebido

Data	Model	Classif.	L_{MSE}			$L_{MSE-log}$		
			Acc	Sen	F1	Acc	Sen	F1
VGG19	ASPD-1BiRe	LR	0.923	0.9	0.927	0.903	0.9	0.915
		RF	0,783	0,65	0,756	0,803	0,683	0,777
		SV	0,787	0,667	0,765	0,787	0,667	0,765
	ASPD-2BiRe	LR	0,903	0,9	0,915	0,883	0,783	0,873
		RF	0,77	0,667	0,757	0,81	0,717	0,798
		SV	0,807	0,7	0,794	0,767	0,617	0,734
	ASPD-3BiRe	LR	0,85	0,783	0,838	0,87	0,75	0,847
		RF	0,807	0,683	0,78	0,787	0,667	0,762
		SV	0,83	0,733	0,81	0,753	0,617	0,714

Tabla 3. La tabla muestra las métricas correspondientes de clasificación utilizando las configuraciones planteadas de modelo (Model), clasificador (Classif.) y la función de pérdida usada para el conjunto de datos proveniente de la VGG19.

Data	Model	Classif.	L_{MSE}			$L_{MSE-log}$		
			Acc	Sen	F1	Acc	Sen	F1
Xception	ASPD-1BiRe	LR	0,86	0,81	0,865	0,803	0,75	0,798
		RF	0,73	0,517	0,66	0,727	0,467	0,627
		SV	0,707	0,467	0,616	0,727	0,517	0,654
	ASPD-2BiRe	LR	0,88	0,817	0,877	0.86	0.817	0.865
		RF	0,827	0,667	0,787	0,767	0,583	0,714
		SV	0,747	0,6	0,707	0,73	0,55	0,681
	ASPD-3BiRe	LR	0,827	0,683	0,791	0,763	0,583	0,698
		RF	0.883	0.783	0.873	0,8	0,65	0,758
		SV	0,79	0,617	0,738	0,707	0,5179	0,632

Tabla 4. La tabla muestra las métricas correspondientes de clasificación utilizando las configuraciones planteadas de modelo (Model), clasificador (Classif.) y la función de pérdida usada para el conjunto de datos proveniente de la Xception.

generado por el modelo de autoaprendizaje SPD. Además, no se presenta una tendencia por la influencia de la profundidad de los modelos (número de bloques BiRe) en el rendimiento de los clasificadores. Particularmente, para la regresión logística (LR), se observa un deterioro de la capacidad discriminatoria con el aumento de la profundidad; por el contrario, para el bosque de decisión aleatorio (RF) se muestra tendencia a mejorar la capacidad de clasificación con el aumento de profundidad, inclusive llegando a superar a la regresión logística; y para la máquina

de soporte vectorial (SV) no se presentan patrones concluyentes.

Con el fin de comprender mejor las representaciones embebidas que se están clasificando por las dos mejores configuraciones del método, se utilizó el método de análisis de componentes principales (PCA) para reducir la representación a un espacio bidimensional. El resultado se muestra en la figura 6. En la figura, se puede observar que ambos espacios embebidos resultan similares y difíciles de diferenciar a simple vista. Por tanto, para medir la separación espacial entre estos puntos calculamos los centroides de los conjuntos de vectores embebidos de cada clase. Sin embargo, al calcular la distancia euclidiana (*norma l2*) entre los centroides de las clases en los diferentes espacios, se encontró que en el espacio ajustado con MSE, la distancia es de 1,52, mientras que en el espacio ajustado con MSElog, la distancia es de 1,46.

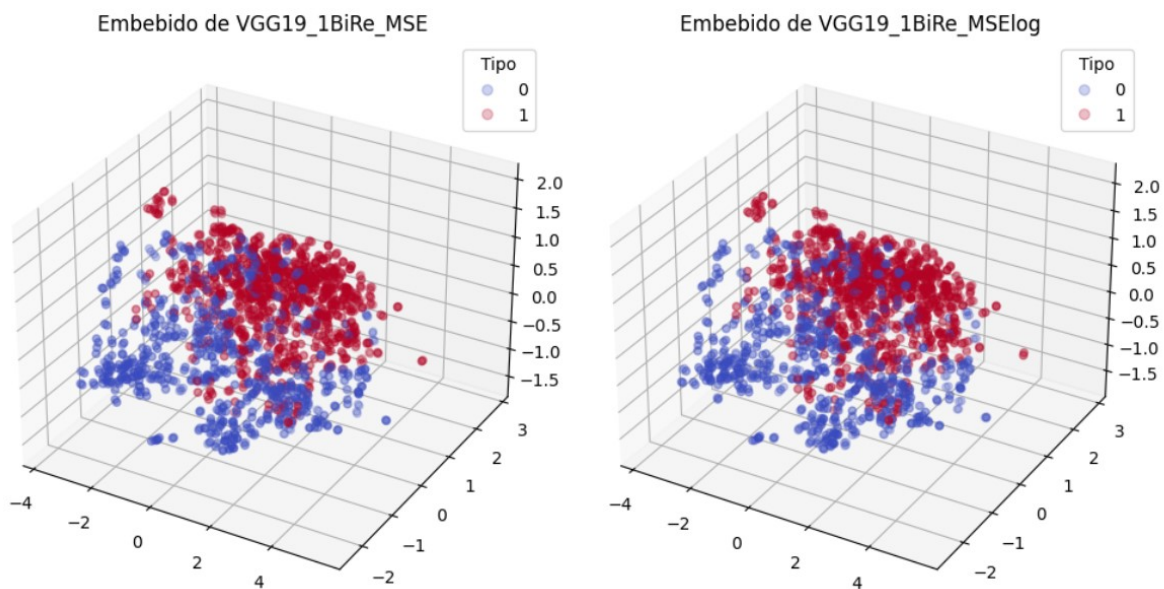


Figura 6. Por medio del análisis de componentes principales (PCA), se redujo a 3 la dimensionalidad de los embebidos usados en la clasificación de los patrones oculomotores .

Además, para evaluar la robustez de las dos mejores configuraciones del método, se calculó la curva ROC para cada una de las correspondientes particiones del k-fold. La Figura 7 presenta estas curvas y su respectiva area bajo la curva ROC (AUC). Con un AUC medio de aproximadamente 0.975 para ambas configuraciones, se demuestra la capacidad de estas configuraciones

para discriminar entre clases de manera efectiva y precisa. Estos resultados refuerzan la robustez del método propuesto, demostrando incluso una capacidad de generalización destacada en diferentes particiones del k-fold.

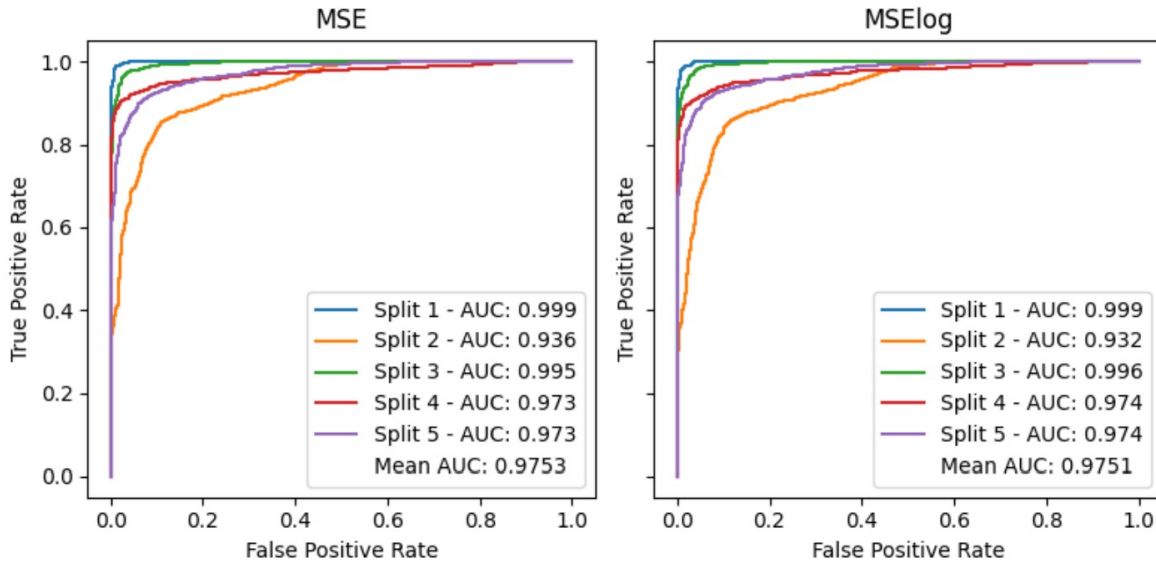


Figura 7. ROC curves of the two best methods with VGG19 data set, ASPD-1BiRe architecture, fitted with L_{MSE} (left) $L_{log-MSE}$ (right) and ranked with logistic regression.

El método propuesto se evaluó con respecto a otras estrategias propuestas en el estado del arte bajo el mismo conjunto de datos. Para esto, el mejor modelo ASPD-1BiRe, con las características de la VGG19, ajustado con L_{MSE} se comparo con el método propuesto por Salazar et. al.¹⁴, el cual propone una covarianza compacta y algoritmos de aprendizaje automático para la clasificación. Se consideró también el modelo convolucional SPD (ConvSPD) que integra redes convolucionales y capas de aprendizaje SPD¹⁶. mismo conjunto de datos planteado (Cuadro 5). En la Tabla 5) se presenta la comparación en métricas de clasificación con los métodos mencionados. En estos resultados se evidencia nuestro avance respecto al enfoque de clasificación de matrices de covarianza que también son reducidas mediante un enfoque no supervisado. Además, se puede observar que, el enfoque ConvSPD que supera nuestros resultados, también considera aspectos geométricos. Sin embargo, nuestro método se posiciona de manera competitiva al aprender representaciones compactas de forma no supervisada, permitiendo una mayor

flexibilidad durante el proceso de entrenamiento.

método	AUC	Acc	Sen	F1
Covarianza + ML	0.88	0.808	-	-
Conv-Riemanniano supervisado	-	0.978	0.972	0.978
Método propuesto	0.97	0.92	0.90	0.93

Tabla 5. Esta tabla recopila información de las métricas obtenidas por dos enfoques propuestos en el estado del arte Covarianza + ML⁵³ y Conv-Riemanniano supervisado⁵⁴.

7. CONCLUSIONES Y TRABAJO FUTURO

En este trabajo se presenta una estrategia de aprendizaje no supervisado para caracterizar representaciones convolucionales en descriptores Riemannianos compactos. Las configuraciones utilizadas con el método propuesto demostraron una convergencia notable, especialmente cuando se emplea la función de pérdida L_{MSE} . Se validó su capacidad para discriminar patrones oculomotores relacionados con el Parkinson, logrando una precisión promedio del 92,3% y un AUC promedio del 97,4% en su mejor configuración. Esta configuración presenta una mejora significativa en comparación con otro método que también genera representaciones compactas mediante análisis de componentes principales (consulte la Tabla 5). Dado que el aprendizaje de las representaciones se realiza de forma no supervisada, este enfoque se posiciona de manera competitiva como método para generar descriptores compactos.

También se mostró que el aumento en la profundidad de la arquitectura no se traduce en una mejora significativa en el espacio embebido. Esto se refiere a la capacidad de clasificar los patrones de fijación o reconstruir las representaciones originales, lo cual, es de esperarse en un proceso de codificación-decodificación, ya que una menor dimensionalidad implica inevitablemente una pérdida de información.

En particular, se pudo observar un deterioro en la capacidad discriminatoria de la regresión logística (LR) a medida que aumenta la profundidad, mostrando que las representaciones embebidas resultantes de arquitecturas con mayor profundidad son menos separables por un hiperplano. Por otro lado, en el caso del bosque de decisión aleatorio (RF), se observó una tendencia a mejorar la capacidad de clasificación a medida que aumenta la profundidad, e incluso puede llegando a superar a la regresión logística en una configuración, mostrando que el aumento de la profundidad tiende a formar grupos separables por varios umbrales en lugar de un solo hiperplano.

Además, se observó la influencia de las matrices de entrada a la arquitectura (X) tanto en el proceso de reconstrucción como de clasificación. Dado que el conjunto de datos que posee

una gran parte de sus autovalores siendo regularizados presenta unas métricas inferiores en la mayoría de los clasificadores, creemos que la generación de este conjunto es lo que está restringiendo la calidad de las representaciones Riemannianas profundas. Por tanto, concluimos que la generación de un conjunto apropiado de datos parece ser un factor clave para generar mejores representaciones Riemannianas embebidas.

Específicamente se ha identificado un enfoque que demuestra un mejor rendimiento en clasificación, al conectar capas convolucionales preentrenadas a un codificador Riemanniano para la clasificación de patrones oculomotores⁵⁵, sin embargo, utiliza un aprendizaje supervisado. Entonces, estas representaciones minimizan y ajustan las matrices SPD con respecto a etiquetas brindadas por expertos, lo cual puede sesgar la representación e inducir a errores en su predicción. Además, el uso de esquemas supervisados refleja una limitación para la obtención de datos y generalización de las arquitecturas. Por ello, como trabajo futuro, se plantea la inclusión de capas convolucionales entrenables que permitan ajustar los mapas de activación y mejorar las representaciones compactas en los descriptores Riemannianos que utiliza inicialmente la arquitectura codificador-decodificador, manteniendo un aprendizaje auto-supervisado.

⁵⁵ Juan Olmos, Antoine Manzanera y Fabio Martínez. “An Oculomotor Digital Parkinson Biomarker from a Deep Riemannian Representation”. En: *Pattern Recognition and Artificial Intelligence: Third International Conference, ICPRAI 2022, Paris, France, June 1–3, 2022, Proceedings, Part I*. Springer. 2022, págs. 677-687.

BIBLIOGRAFIA

- Abdulhay, Enas, N Arunkumar, Kumaravelu Narasimhan et al. “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. 23).
- Absil, P-A, Robert Mahony y Rodolphe Sepulchre. *Optimization algorithms on matrix manifolds*. Princeton University Press, 2009 (vid. pág. 37).
- Acharya, Dinesh et al. “Covariance pooling for facial expression recognition”. En: *Proceedings of the IEEE conference on computer vision and pattern recognition workshops*. 2018, págs. 367-374 (vid. pág. 12).
- Ajay, Jerry, Chen Song, Aosen Wang et al. “A pervasive and sensor-free Deep Learning system for Parkinsonian gait analysis”. En: *2018 IEEE EMBS International Conference on Biomedical Health Informatics (BHI)*. 2018, págs. 108-111 (vid. pág. 23).
- Antoniades, CA y C Kennard. “Ocular motor abnormalities in neurodegenerative disorders”. En: *Eye* 29.2 (2015), pág. 200 (vid. pág. 14).
- Armstrong, R. A. “Oculo-Visual Dysfunction in Parkinson’s Disease”. En: *Journal of Parkinson’s Disease* 5 (2015). 4, págs. 715-726 (vid. pág. 16).
- Arsigny, Vincent et al. “Geometric means in a novel vector space structure on symmetric positive-definite matrices”. En: *SIAM journal on matrix analysis and applications* 29.1 (2007), págs. 328-347 (vid. pág. 31).

- Arsigny, Vincent et al. “Geometric Means in a Novel Vector Space Structure on Symmetric Positive-Definite Matrices”. En: *SIAM Journal on Matrix Analysis and Applications* 29.1 (2007), págs. 328-347 (vid. pág. 20).
- Baldi, Pierre. “Autoencoders, unsupervised learning, and deep architectures”. En: *Proceedings of ICML workshop on unsupervised and transfer learning*. JMLR Workshop y Conference Proceedings. 2012, págs. 37-49 (vid. pág. 18).
- Błaszczyk, JW. “Motor deficiency in Parkinson’s disease”. En: *Acta neurobiologiae experimentalis* 58.1 (1998), 79—93 (vid. pág. 14).
- Caramia, Carlotta, Diego Torricelli, Maurizio Schmid et al. “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. 23).
- Carson, Tana B y Stephen Z Sutton. *Application for smart phone or related devices for use in assessment of vestibulo-ocular reflex function*. US Patent App. 15/569,472. 2018 (vid. pág. 11).
- Chollet, François. “Xception: Deep learning with depthwise separable convolutions”. En: *Proceedings of the IEEE conference on computer vision and pattern recognition*. 2017, págs. 1251-1258 (vid. pág. 41).
- Dong, Ganggang y Gangyao Kuang. “Target Recognition in SAR Images via Classification on Riemannian Manifolds”. En: *IEEE Geoscience and Remote Sensing Letters* 12.1 (2015), págs. 199-203 (vid. pág. 22).

- Duval, Christian y Anne Beuter. “Fluctuations in Tremor at Rest and Eye Movements during Ocular Fixation in Subjects with Parkinson’s Disease”. En: *Parkinsonism & Related Disorders* 4 (1998), págs. 91-97 (vid. pág. 15).
- Ekker, Merel S, Sabine Janssen, Klaus Seppi et al. “Ocular and visual disorders in Parkinson’s disease: common but frequently overlooked”. En: *Parkinsonism & related disorders* 40 (2017), págs. 1-10 (vid. págs. 11, 14, 21, 23).
- Ezdon, Cindy. “Principles of Neural Science”. En: *The American Journal of Occupational Therapy* 41.1 (1987), págs. 59-59 (vid. págs. 14, 23).
- Feigin, Valery L., Emma Nichols, Tahiya Alam et al. “Global, regional, and national burden of neurological disorders, 1990–2016: a systematic analysis for the Global Burden of Disease Study 2016”. En: *The Lancet Neurology* 18.5 (2019), págs. 459-480 (vid. págs. 10, 14, 23).
- Fukushima, Kikuro, Norie Ito, Graham R Barnes et al. “Impaired smooth-pursuit in Parkinson’s disease: normal cue-information memory, but dysfunction of extra-retinal mechanisms for pursuit preparation and execution”. En: *Physiological reports* 3.3 (2015), e12361 (vid. pág. 16).
- Gitchel, George T, Paul A Wetzal y Mark S Baron. “Pervasive ocular tremor in patients with Parkinson disease”. En: *Archives of neurology* 69.8 (2012), págs. 1011-1017 (vid. págs. 11, 15, 21, 24).
- Hu, Kun et al. “Vision-Based Freezing of Gait Detection with Anatomic Patch Based Representation”. En: *Computer Vision – ACCV 2018*. Ed. por C. V. Jawahar et al. Cham: Springer International Publishing, 2019, págs. 564-576 (vid. pág. 23).
- Huang, Zhiwu y Luc Van Gool. “A riemannian network for spd matrix learning”. En: *Thirty-first AAAI conference on artificial intelligence*. 2017 (vid. págs. 12, 21, 22, 30, 32, 36, 37).

- Huang, Zhiwu y Luc Van Gool. “A riemannian network for spd matrix learning”. En: *Proceedings of the AAAI conference on artificial intelligence*. Vol. 31. 1. 2017 (vid. pág. 31).
- Huang, Zhiwu et al. “Projection metric learning on Grassmann manifold with application to video based face recognition”. En: *Proceedings of the IEEE conference on computer vision and pattern recognition*. 2015, págs. 140-149 (vid. pág. 21).
- Jankovic, Joseph. “Parkinson’s disease: clinical features and diagnosis”. En: *Journal of neurology, neurosurgery & psychiatry* 79.4 (2008), págs. 368-376 (vid. pág. 10).
- Krauzlis, Richard J, Laurent Goffart y Ziad M Hafed. “Neuronal control of fixation and fixational eye movements”. En: *Philosophical Transactions of the Royal Society B: Biological Sciences* 372.1718 (2017) (vid. pág. 15).
- Krizhevsky, Alex, Ilya Sutskever y Geoffrey E Hinton. “Imagenet classification with deep convolutional neural networks”. En: *Communications of the ACM* 60.6 (2017), págs. 84-90 (vid. pág. 40).
- Kulis, Brian, Mátyás A Sustik e Inderjit S Dhillon. “Low-Rank Kernel Learning with Bregman Matrix Divergences.” En: *Journal of Machine Learning Research* 10.2 (2009) (vid. pág. 21).
- Lai, Hsin-Yu, Gladynel Saavedra-Peña, Charles Sodini et al. “Enabling Saccade Latency Measurements with Consumer-Grade Cameras”. En: *2018 25th IEEE International Conference on Image Processing (ICIP)*. 2018, págs. 3169-3173 (vid. págs. 11, 21).
- Lal, Vivek y Daniel Truong. “Eye movement abnormalities in movement disorders”. En: *Clinical Parkinsonism & Related Disorders* 1 (2019), págs. 54-63 (vid. págs. 11, 14, 21, 23).

- Larrazabal, Agostina J, CE García Cena y César Ernesto Martínez. “Video-oculography eye tracking towards clinical applications: A review”. En: *Computers in biology and medicine* 108 (2019), págs. 57-66 (vid. págs. 11, 21).
- LaValley, Michael P. “Logistic regression”. En: *Circulation* 117.18 (2008), págs. 2395-2399 (vid. pág. 39).
- Lewandowski, Tomasz y Anna Lewandowska. “Computer-aided processing of the oculomotor signal”. En: *ITM Web of Conferences*. Vol. 21. EDP Sciences. 2018, pág. 00025 (vid. pág. 24).
- Litjens, Geert, Thijs Kooi, Babak Ehteshami Bejnordi et al. “A survey on deep learning in medical image analysis”. En: *Medical Image Analysis* 42 (2017), págs. 60-88 (vid. pág. 16).
- MacAskill, Michael R. y Tim J. Anderson. “Eye movements in neurodegenerative diseases”. En: *Current Opinion in Neurology* 29.1 (2016) (vid. págs. 14, 23).
- Minh, Hà Quang y Vittorio Murino. *Covariances in computer vision and machine learning*. Springer, 2018, págs. 7-17 (vid. pág. 19).
- Olmos, Juan, Antoine Manzanera y Fabio Martínez. “An Oculomotor Digital Parkinson Biomarker from a Deep Riemannian Representation”. En: *Pattern Recognition and Artificial Intelligence: Third International Conference, ICPRAI 2022, Paris, France, June 1–3, 2022, Proceedings, Part I*. Springer. 2022, págs. 677-687 (vid. págs. 12, 21, 22, 24, 48, 51).
- Pal, Mahesh. “Random forest classifier for remote sensing classification”. En: *International journal of remote sensing* 26.1 (2005), págs. 217-222 (vid. pág. 39).
- Poewe, Werner, Klaus Seppi, Caroline M. Tanner et al. “Parkinson disease”. En: *Nature Reviews Disease Primers* 3 (2017), págs. 1-21 (vid. págs. 10, 14, 22, 23).

- Przybyszewski, Andrzej W, Stanislaw Szlufik et al. “Machine learning on the video basis of slow pursuit eye movements can predict symptom development in Parkinson’s patients”. En: *Intelligent Information and Database Systems: 7th Asian Conference, ACIIDS 2015, Bali, Indonesia, March 23-25, 2015, Proceedings, Part II 7*. Springer International Publishing. 2015, págs. 268-276 (vid. pág. 11).
- Salazar, Isail et al. “A convolutional oculomotor representation to model parkinsonian fixational patterns from magnified videos”. En: *Pattern Analysis and Applications 24* (2021), págs. 445-457 (vid. págs. 12, 21, 22, 24, 48).
- Shibasaki, Hiroshi, Sadatoshi Tsuji y Yoshigoro Kuroiwa. “Oculomotor Abnormalities in Parkinson’s Disease”. En: *Archives of Neurology* 36.6 (jun. de 1979), págs. 360-364 (vid. pág. 16).
- Simonyan, Karen y Andrew Zisserman. “Very deep convolutional networks for large-scale image recognition”. En: *arXiv preprint arXiv:1409.1556* (2014) (vid. pág. 40).
- Turcano, Pierpaolo et al. “Early ophthalmologic features of Parkinson’s disease: a review of preceding clinical and diagnostic markers”. En: *Journal of neurology* 266 (2019), págs. 2103-2111 (vid. págs. 11, 14, 21, 23).
- Tuzel, Oncel, Fatih Murat Porikli y Peter Meer. “Pedestrian Detection via Classification on Riemannian Manifolds”. En: *IEEE Transactions on Pattern Analysis and Machine Intelligence* 30 (2008), págs. 1713-1727 (vid. págs. 19, 22).
- Vemuri, Vijay K. “The Hundred-Page Machine Learning Book”. En: *Journal of Information Technology Case and Application Research* 22.2 (2020), págs. 136-138 (vid. pág. 16).
- Wang, Lipo. *Support vector machines: theory and applications*. Vol. 177. Springer Science & Business Media, 2005 (vid. pág. 39).

- Wang, Wu et al. “Early detection of Parkinson’s disease using deep learning and machine learning”. En: *IEEE Access* 8 (2020), págs. 147635-147646 (vid. págs. 11, 12).
- Weil, Rimona S., Anette E. Schrag, Jason D. Warren et al. “Visual dysfunction in Parkinson’s disease”. En: *Brain* 139.11 (2016), págs. 2827-2843 (vid. págs. 10, 11, 15).
- Ying, Xue. “An overview of overfitting and its solutions”. En: *Journal of physics: Conference series*. Vol. 1168. IOP Publishing. 2019, págs. 022022 (vid. págs. 17).
- Zhai, Shuangfei y Zhongfei Zhang. “Semisupervised autoencoder for sentiment analysis”. En: *Proceedings of the AAAI Conference on Artificial Intelligence*. Vol. 30. 1. 2016 (vid. págs. 34).
- Zhou, Zhi-Hua. “A brief introduction to weakly supervised learning”. En: *National science review* 5.1 (2018), págs. 44-53 (vid. págs. 17).