

**CLASIFICACIÓN DE LESIONES DE CÁNCER DE PRÓSTATA
CARACTERIZANDO ESPACIOS EMBEBIDOS DE OBSERVACIONES MRI**

CÉSAR LUIS HURTADO RODRÍGUEZ

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

2024

**CLASIFICACIÓN DE LESIONES DE CÁNCER DE PRÓSTATA
CARACTERIZANDO ESPACIOS EMBEBIDOS DE OBSERVACIONES MRI**

CÉSAR LUIS HURTADO RODRÍGUEZ

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

Director:

**Fabio Martínez Carrillo
Ph.D**

Codirector:

**Juan Andrés Olmos Rojas
MS.c**

Tutor:

**Yesid Alfonso Gutiérrez Guate
MS.c**

**UNIVERSIDAD INDUSTRIAL DE SANTANDER
FACULTAD DE INGENIERÍAS FISICOMECAICAS
ESCUELA DE INGENIERÍA DE SISTEMAS E INFORMÁTICA
PROGRAMA DE PREGRADO EN INGENIERÍA DE SISTEMAS E
INFORMÁTICA
BUCARAMANGA**

2024

AGRADECIMIENTOS

Primeramente, gracias a Dios por permitirme vivir, por darme fortaleza cada que la necesitaba para poder lograr este objetivo. Agradecerle a quienes admiro mucho, quienes fueron mis guías durante este proceso, el profesor Fabio Martínez y Juan Olmos. Al profesor Fabio agradecerle por su arduo trabajo como investigador, por su excelente guía, por su gran paciencia, disposición y compromiso durante el desarrollo de este trabajo, llegando a ser uno de los investigadores que mas admiro en la universidad. También a Olmos, quien fue mi codirector, un matemático teso que me brindó siempre su apoyo académico, estando siempre a disposición, guiándome en cada momento, aclarándome cualquier duda por minúscula que fuese. En general, gracias a la guía y acompañamiento de estos grandes investigadores, por influir positivamente en mí, enseñándome que siempre se puede dar más. Igualmente, al grupo de investigación BIVL2AB por brindarme conocimiento por medio de sus excelentes integrantes. También quiero agradecer a quienes son mis ídolos número uno en la vida, mi Madre y mi Padre. A mi Madre, una mujer que admiro mucho, gracias por siempre estar conmigo apoyándome en todo mal o buen momento, por su comprensión y paciencia, por siempre querer lo mejor para mí y para mi hermano luchando toda su vida por nosotros. A mi Padre, que es el hombre más correcto que conozco, agradecerle por su apoyo, por su disposición ante cualquier adversidad, por siempre hablarme con sinceridad, gracias por ser el ejemplo que siempre voy a seguir. Por ellos es que soy quien soy y estoy donde estoy, por lo que mi principal objetivo en la vida es velar por el bienestar de ellos. También agradecer a mi pareja por siempre estar para mí, por apoyarme incondicionalmente en mis peores momentos, por amarme tan bonito, por siempre darme ánimos para seguir y nunca bajarme del barco, por escucharme cuando más quería expresarme, por último, por estar conmigo en momentos en los que nadie más estaría. Además, agradecer a mis compañeros y amigos de carrera, con los que compartí los mejores momentos de esta etapa de mi vida, con los que inicié y aun sigue esa amistad. Por ultimo a los "Kunis" por brindarme su amistad sincera y por hacerme pasar los mejores momentos con ellos. Por otra parte, agradecer a Emmanuel Gazmey

Santiago y Agustín Carlos Roberto García, que son en quienes me apoyé en cada mal momento que pasaba, brindándome una tranquilidad inefable.

Por último, el resultado de este trabajo es gracias al esfuerzo que le brindé, porque a pesar de las adversidades que se me presentaban, nunca desistí. El resultado de este trabajo fue porque lo trabajé y le metí la pasión con el talento necesario. Mi futuro siempre lo estoy buscando, nunca lo dejé al azar.

CONTENIDO

	page
INTRODUCCIÓN	12
1. FUNDAMENTOS Y TRABAJO PREVIO	16
1.1. CÁNCER DE PRÓSTATA Y SECUENCIAS BP-MRI	16
1.1.1 Imagen ponderada en T2 (T2W).	19
1.1.2 Imagen ponderada en difusión (DWI) y mapas ADC.	20
1.2. REPRESENTACIONES EMBEBIDAS	24
1.2.1 Aprendizaje Contrastivo Supervisado (SLC).	24
1.2.2 Representaciones geométricas SPD.	27
1.3. SISTEMAS DE DIAGNÓSTICO ASISTIDO POR COMPUTADOR	28
2. PROBLEMA DE INVESTIGACIÓN	32
3. OBJETIVOS	33
4. MÉTODO PROPUESTO	34
4.1. ESQUEMA CONVOLUCIONAL 3D PARA LA EXTRACCIÓN DE CARACTERÍSTICAS MRI	35
4.2. ESQUEMA DE APRENDIZAJE CONTRASTIVO SUPERVISADO (SCL)	36
4.3. CLASIFICACIÓN DE EMBEBIDOS (<i>CONTRASTIVO</i>).	39
4.4. EMBEBIDOS SPD GEOMÉTRICOS	41
4.4.1 Clasificación de embebidos geométricos SPD proyectados en el espacio tangente ($\text{Log}(SPD)$).	42
4.4.2 Clasificación mediante redes profundas SPD (<i>SPDnet</i>).	43

5. DISEÑO EXPERIMENTAL	46
5.1. CONJUNTO DE DATOS	46
5.2. CONFIGURACIÓN EXPERIMENTAL	48
5.2.1 Arquitectura red encoder y proyección.	48
5.2.2 Esquema de aprendizaje contrastivo supervisado (SCL).	49
5.2.3 Embebidos SPD.	50
5.2.4 Arquitectura SPDnet.	50
5.2.5 Clasificación de embebidos.	50
5.3. MÉTRICAS	53
6. EVALUACIÓN Y RESULTADOS	54
7. CONCLUSIONES Y TRABAJO FUTURO	62
BIBLIOGRAFÍA	65

LISTA DE FIGURAS

	page
Figura 1. Ejemplos de hallazgos en secuencias T2W	20
Figura 2. Ejemplos de hallazgos en secuencias DWI	22
Figura 3. Ejemplos de hallazgos en secuencias ADC	23
Figura 4. Esquema de aprendizaje contrastivo	25
Figura 5. Pipeline del método propuesto	35
Figura 6. Minado de datos	38
Figura 7. Mapeo Riemanniano de la función logaritmo	43
Figura 8. Parches volumétricos bp-MRI	48
Figura 9. Métodos de clasificación de embebidos	52
Figura 10. Rendimiento de los modelos en escenarios con pocos datos	57
Figura 11. Distribución de predicciones en grados ISUP	60

LISTA DE TABLAS

	page
Tabla 1. Cantidad de estudios	47
Tabla 2. Variables clínicas del conjunto de datos	47
Tabla 3. Resultados de los métodos propuestos	54

RESUMEN

TÍTULO: CLASIFICACIÓN DE LESIONES DE CÁNCER DE PRÓSTATA CARACTERIZANDO ESPACIOS EMBEBIDOS DE OBSERVACIONES MRI. *

AUTOR: CÉSAR LUIS HURTADO RODRÍGUEZ **

PALABRAS CLAVE: CÁNCER DE PRÓSTATA, LESIÓN CSPCA, BP-MRI, EMBEBIDOS.

DESCRIPCIÓN: El cáncer de próstata es el cáncer más frecuente en los hombres. A nivel mundial, en el 2022, más de 1.4 millones de personas fueron diagnosticadas y más de 397 mil decesos fueron asociados a esta enfermedad. En la rutina clínica, las técnicas comúnmente utilizadas para el diagnóstico son el test de Antígeno Prostático Específico (PSA) y el examen digital rectal (DRE). Sin embargo, estos métodos son invasivos y tienen baja especificidad. Actualmente, el análisis de resonancias magnéticas (MRI) es una alternativa con mayor capacidad de detección de lesiones de cáncer de próstata. Sin embargo, su análisis depende de la experiencia del lector. Recientemente, métodos de aprendizaje profundo han emergido como una alternativa para la clasificación de lesiones clínicamente significativas (csPCa) en MRI. Sin embargo, debido a la alta variabilidad intra- e inter-MRI y a la limitación de datos, existe una brecha en la tarea de clasificar estas lesiones. En este trabajo se desarrolló un enfoque de aprendizaje contrastivo para la clasificación de lesiones csPCa en bp-MRI. Inicialmente, bajo un marco de aprendizaje contrastivo, se ajustó una red convolucional dedicada a aprender características profundas mediante la proyección en vectores de baja dimensión. Particularmente, este propone la codificación de matrices simétricas definidas positivas (SPD) a partir de las características profundas convolucionales 3D aprendidas en la tarea contrastiva para la construcción de un espacio embebido geométrico. Para la clasificación de estas matrices SPD, se utilizaron modelos de aprendizaje de máquina y modelos de aprendizaje profundo geométrico que consideran la geometría Riemanniana de las matrices SPD. Los métodos implementados se validaron sobre un conjunto de datos público respecto a la tarea de clasificación de lesiones csPCa, el método propuesto alcanzó un AUC-ROC del 0.93, superando a la discriminación desde los vectores contrastivos con 0.85. Adicionalmente, al reducir los datos de entrenamiento a un 20%, el método propuesto demostró ser robusto al alcanzar un AUC-ROC de 0.91, mientras que para el baseline, esta métrica se redujo al 0.77.

* Trabajo de investigación

** Facultad de Ingenierías Físicomecánicas. Escuela de Ingeniería de Sistemas e Informática. Director: Fabio Martínez, Ph.D. Codirector: Juan Andrés Olmos Rojas, M.Sc.

ABSTRACT

TITLE: CLASSIFICATION OF PROSTATE CANCER LESIONS BY CHARACTERIZING EMBEDDED SPACES OF MRI OBSERVATIONS. *

AUTHOR: CESAR LUIS HURTADO RODRIGUEZ **

KEYWORDS: PROSTATE CANCER, CSPCA LESION, BP-MRI, EMBEDDINGS.

DESCRIPTION: Prostate cancer is the most common cancer in men. Worldwide, in 2022, over 1.4 million people were diagnosed, and more than 397,000 deaths were associated with this disease. In clinical practice, commonly used diagnostic techniques are the Prostate-Specific Antigen (PSA) test and digital rectal examination (DRE). However, these methods are invasive and have low specificity. Currently, magnetic resonance imaging (MRI) analysis is an alternative with a higher capacity for detecting prostate cancer lesions. However, its interpretation depends on the reader's experience. Recently, deep learning methods have emerged as an alternative for classifying clinically significant lesions (csPCa) in MRI. However, due to high intra- and inter-MRI variability and data limitations, there remains a gap in the task of classifying these lesions. In this work, a contrastive learning approach was developed for the classification of csPCa lesions in bp-MRI. Initially, within a contrastive learning framework, a convolutional network was tuned to learn deep features through projection into low-dimensional vectors. Specifically, it proposes encoding symmetric positive definite (SPD) matrices from the deep 3D convolutional features learned in the contrastive task to construct a geometric embedding space. For the classification of these SPD matrices, machine learning models and geometric deep learning models were used, considering the Riemannian geometry of SPD matrices. The implemented methods were validated on a public dataset for the csPCa lesion classification task, with the proposed method achieving an AUC-ROC of 0.93, outperforming discrimination from contrastive vectors, which scored 0.85. Additionally, when reducing the training data to 20%, the proposed method proved robust, achieving an AUC-ROC of 0.91, whereas for the baseline, this metric dropped to 0.77.

* Research work

** Faculty of Physics-Mechanics Engineering. School of Systems Engineering and Informatics. Advisor: Fabio Martínez, Ph.D. Co-advisor: Juan Andrés Olmos Rojas, M.Sc.

INTRODUCCIÓN

El cáncer de próstata (PCa) es el cuarto tipo de cáncer más diagnosticado a nivel mundial y ocupa el segundo lugar como cáncer más frecuente en hombres ¹. En 2022 se registraron más de 1.4 millones de nuevos casos y más de 397.000 decesos a nivel mundial, incluso se estima que para el año 2040 las cifras de mortalidad se dupliquen. Particularmente, en Colombia se contabilizaron 4.304 casos nuevos, siendo la segunda causa principal de muertes por cáncer en hombres con más de 2.000 fallecimientos por año según cifras del año 2022, de la Agencia Internacional para la Investigación sobre el Cáncer ¹². En la rutina clínica, un diagnóstico temprano permite soportar la identificación de lesiones de PCa de alto riesgo, permitiendo un tratamiento oportuno de la enfermedad ³. Para apoyar este diagnóstico, existen diversos métodos como el examen sanguíneo para el análisis del antígeno prostático (PSA, por sus siglas en inglés). A pesar de ser uno de los más utilizados en la rutina clínica, se ha reportado una baja especificidad de alrededor de un 36% ⁴. Esto se traduce en altas tasas de falsos positivos, lo que lleva a procedimientos innecesarios y sobre diagnósticos ⁵. Otro método utilizado en la rutina clínica es el examen digital rectal (DRE, por sus siglas en inglés), el cual ha demostrado no reducir significativamente la mortalidad y también conlleva a un alto número de falsos positivos

¹ Freddie BRAY et al. “Global cancer statistics 2022: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries”. In: *CA: a cancer journal for clinicians* 74.3 (2024), pp. 229–263.

² Jacques FERLAY et al. “Estimating the global cancer incidence and mortality in 2018: GLOBOCAN sources and methods”. In: *International journal of cancer* 144.8 (2019), pp. 1941–1953.

³ Jack CUZICK et al. “Prevention and early detection of prostate cancer”. In: *The lancet oncology* 15.11 (2014), e484–e492.

⁴ Leen NAJI et al. “Digital rectal examination for prostate cancer screening in primary care: a systematic review and meta-analysis”. In: *The Annals of Family Medicine* 16.2 (2018), pp. 149–154.

⁵ Gillian MURPHY et al. “The expanding role of MRI in prostate cancer”. In: *AJR Am J Roentgenol* 201.6 (2013), pp. 1229–1238.

⁴. Además, este método es altamente subjetivo a la experiencia del experto y puede omitir lesiones ubicadas en zonas lejanas a la pared rectal ⁴.

Actualmente, las imágenes de resonancia magnética (MRI, por sus siglas en inglés) se han convertido en una herramienta útil para el diagnóstico en etapas previas a biopsia. Se ha demostrado que con la resonancia se puede aumentar la precisión en la detección, localización y estratificación del tumor ⁵. En la práctica clínica, comúnmente se utiliza la resonancia multiparamétrica (mp-MRI), la cual implementa sustancias de contraste, permitiendo determinar las propiedades capilares, morfológicas y vasculares de los tejidos de la glándula prostática, sin embargo, el uso de estos agentes de contraste pueden generar efectos secundarios en el paciente como por ejemplo alergias o incluso fibrosis sistémica nefrogénica ⁶. Además, el tiempo del proceso de adquisición de estas imágenes puede ser más prolongado que el de otras secuencias de resonancia magnética, ya que implica la captura de múltiples imágenes a lo largo del tiempo para evaluar la perfusión de contraste en el tejido ⁷. En su lugar, se ha optado por el uso de la resonancia magnética biparamétrica (bp-MRI), que no considera estas secuencias de contraste. Se ha demostrado que el uso de bp-MRI alcanza una detección de PCa equiparable con lo logrado por parte de la mp-MRI tradicional, eliminando los posibles riesgos asociados con el gadolinio y reduciendo costos, tiempo y exposición al contraste ⁸. Estos hallazgos han posicionado a la bp-MRI como una alternativa eficiente y prometedora para soportar la detección de PCa de manera oportuna y no invasiva ⁸. En la rutina clínica, a pesar del aumento en el uso de la bp-MRI para el diagnóstico de PCa, el análisis de estas imágenes representa una labor compleja para los expertos, debido a la variabilidad en apariencia que hay en lesiones de

⁶ Cher Heng TAN et al. “Dynamic contrast-enhanced MRI for the detection of prostate cancer: meta-analysis”. In: *American Journal of Roentgenology* 204.4 (2015), W439–W448.

⁷ Rose M BERMAN et al. “DCE MRI of prostate cancer”. In: *Abdominal radiology* 41 (2016), pp. 844–853.

⁸ Rachael L SHERRER et al. “Comparison of biparametric MRI to full multiparametric MRI for detection of clinically significant prostate cancer”. In: *Prostate cancer and prostatic diseases* 22.2 (2019), pp. 331–336.

cáncer de próstata ⁹. Además, en la lectura de una MRI persiste una variabilidad dependiente al grado de experiencia del médico que realiza el análisis, donde más años de experiencia lleva a mejoras en la sensibilidad, el valor predictivo negativo y la confianza del lector ¹⁰.

Actualmente, el uso de estrategias computacionales ha demostrado su capacidad para clasificar lesiones clínicamente significativas de cáncer de próstata (csPCa), resultando en potenciales herramientas para soportar al experto en el diagnóstico ¹¹. Entre las estrategias existentes, los enfoques que involucran aprendizaje profundo resaltan por una precisión prometedora, que incluso puede asemejarse a la de expertos radiólogos ¹². No obstante, estos métodos requieren de grandes cantidades de datos etiquetados para el correcto aprendizaje de patrones característicos de las lesiones de PCa. Sin embargo, la disponibilidad de datos sigue siendo un obstáculo significativo, especialmente en el ámbito médico, donde la adquisición de datos puede ser costosa y estar sujeta a regulaciones de privacidad ¹³. Ante esta problemática, gran parte de los métodos se han limitado a utilizar estrategias como lo es el aumento excesivo de datos, resultando en un sobre ajuste en el aprendizaje e ignorando la alta variabilidad que existe entre muestras de una misma clase, lo que puede llevar a un bajo rendimiento en el desempeño sobre nuevos conjuntos de datos.

Este trabajo desarrolló y exploró un enfoque de aprendizaje profundo contrastivo para

⁹ Jonathan SHOAG and Christopher E BARBIERI. “Clinical variability and molecular heterogeneity in prostate cancer”. In: *Asian journal of andrology* 18.4 (2016), p. 543.

¹⁰ Bassel R SALKA et al. “Effect of prostate MRI interpretation experience on PPV using PI-RADS version 2: a 6-year assessment among eight fellowship-trained radiologists”. In: *AJR. American journal of roentgenology* 219.3 (2022), p. 453.

¹¹ Olivier ROUVIÈRE et al. “Artificial intelligence algorithms aimed at characterizing or detecting prostate cancer on MRI: How accurate are they when tested on independent cohorts?—a systematic review”. In: *Diagnostic and interventional imaging* 104.5 (2023), pp. 221–234.

¹² Lidia Alcalá MATA et al. “Artificial intelligence–assisted prostate cancer diagnosis: Radiologic–pathologic correlation”. In: *RadioGraphics* 41.6 (2021), pp. 1676–1697.

¹³ Aghiles KEBAILI; Jérôme LAPUYADE-LAHORGUE, and Su RUAN. “Deep learning approaches for data augmentation in medical imaging: a review”. In: *Journal of Imaging* 9.4 (2023), p. 81.

la clasificación de lesiones csPCa a partir de representaciones embebidas de secuencias bp-MRI. En primer lugar, se utilizó una red convolucional 3D para proyectar los hallazgos bp-MRI en vectores de baja dimensión. Posteriormente, se ajustó este extractor de características siguiendo una estrategia de aprendizaje contrastivo en un marco supervisado que considera la alta variabilidad entre muestras de igual clase mientras se aprende a diferenciar información que proviene de muestras de diferente clase. Tradicionalmente, a partir del espacio de vectores embebidos resultante, se realiza la clasificación de estos embebidos, esto es, diferenciar entre proyecciones embebidas de lesiones csPCa y muestras de regiones sanas. En este trabajo se propuso el cálculo de embebidos de segundo orden, considerando matrices simétricas definidas positivas (SPD) que resumen de manera compacta las relaciones en las características profundas extraídas del encoder convolucional 3D. Para la clasificación de estos descriptores SPD, se consideraron modelos clásicos de aprendizaje de máquina y arquitecturas de aprendizaje profundo geométrico que consideran la geometría de la variedad diferenciable Riemanniana a la cual pertenecen estas matrices. El método implementado se validó en un conjunto de datos público con respecto a la tarea de clasificación de lesiones csPCa y su comportamiento en escenarios con pocos datos.

1. FUNDAMENTOS Y TRABAJO PREVIO

1.1. CÁNCER DE PRÓSTATA Y SECUENCIAS BP-MRI

El Cáncer de Próstata (PCa) comienza con el crecimiento anormal de las células en la glándula prostática, lo cual evoluciona en un tejido anormal, comúnmente denominado como tumor. La degeneración celular puede involucrar la activación de oncogenes, lo que transforma a estas células en cancerígenas ¹⁴.

Actualmente, en la rutina clínica, los métodos comúnmente utilizados para el diagnóstico de PCa son: la prueba de Antígeno prostático específico (PSA) y el examen digital rectal (DRE). La prueba de PSA mide la presencia de esta sustancia proteica (Antígeno prostático específico) en la sangre. Aunque el uso de esta prueba como método de tamizaje poblacional demostró una disminución en diferentes barridos poblacionales, trabajos más recientes han revelado que hasta un 50% de hombres con resultados del PSA por encima de la cantidad normal de esta sustancia proteica y que se sometieron a una biopsia, no tenían cáncer, lo cual se traduce a una baja especificidad en el diagnóstico ¹⁵. Por otra parte, el examen DRE consiste en palpar la glándula prostática con el fin de detectar regiones de la próstata que puedan presentar irregularidades, tales como bultos o tejidos consistentes que podrían estar asociados con tumores malignos ¹⁶. A pesar de su utilidad en la práctica para rápidamente caracterizar potencialmente tejidos anormales, esta caracterización está fuertemente ligada a la experiencia del experto, lo que

¹⁴ Amit R PATEL and Eric A KLEIN. “Risk factors for prostate cancer”. In: *Nature clinical practice Urology* 6.2 (2009), pp. 87–95.

¹⁵ Simona FERRARO; Marco BUSSETTI, and Mauro PANTEGHINI. “Serum prostate-specific antigen testing for early detection of prostate cancer: managing the gap between clinical and laboratory practice”. In: *Clinical Chemistry* 67.4 (2021), pp. 602–609.

¹⁶ Onisuru T OKOTIE et al. “Characteristics of prostate cancer detected by digital rectal examination only”. In: *Urology* 70.6 (2007), pp. 1117–1120.

lleva a una subjetividad en el diagnóstico, además de ser un procedimiento invasivo en el que incluso se pueden ignorar lesiones en zonas lejanas o de difícil acceso en la glándula prostática⁴. En etapas más avanzadas de la enfermedad, se suele emplear la biopsia dirigida por ecografía transrectal para estudiar la enfermedad mediante la obtención de pequeñas porciones de tejido, las cuales son posteriormente analizadas en laboratorio mediante imágenes histológicas¹⁷. Sin embargo, durante esta ecografía transrectal existen zonas con apariencia similar a las de una lesión maligna, lo que lleva al análisis de regiones de interés que finalmente no son relacionadas con el cáncer, generando la extracción de tejido innecesario. Además, alrededor del 50% de las lesiones no palpables con un diámetro menor de 1cm no se observan en la ecografía, lo que significa que algunos tejidos cancerígenos pueden ser ignorados. También, el procedimiento de biopsia involucra potenciales complicaciones posteriores a la biopsia como infecciones, riesgo de hospitalización, aumentando las tasas de admisión por problemas urológicos¹⁸¹⁹.

Las imágenes de Resonancia Magnética (MRI) han emergido como una herramienta de apoyo que permite de manera no invasiva visualizar la glándula prostática, así como la apariencia de diferentes regiones de tejido, permitiendo así localizar tumores e incluso determinar su agresividad²⁰. Particularmente, la Resonancia Magnética Multiparamétrica (mp-MRI) es un tipo de imagen ampliamente utilizada en el manejo de pacientes con cáncer de próstata. Estas imágenes se componen de tres secuencias: *T2-weighted imaging* (T2W), *diffusion-weighted imaging* (DWI) y *dynamic contrast-enhanced imaging* (DCE). Estas secuencias brindan informa-

¹⁷ J RAJA et al. “Current status of transrectal ultrasound-guided prostate biopsy in the diagnosis of prostate cancer”. In: *Clinical radiology* 61.2 (2006), pp. 142–153.

¹⁸ Larissa V RODRIGUEZ and Martha K TERRIS. “Risks and complications of transrectal ultrasound guided prostate needle biopsy: a prospective study and review of the literature”. In: *The Journal of urology* 160.6 Part 1 (1998), pp. 2115–2120.

¹⁹ Robert K NAM et al. “Increasing hospital admission rates for urological complications after transrectal ultrasound guided prostate biopsy”. In: *The Journal of urology* 183.3 (2010), pp. 963–969.

²⁰ Amita SHUKLA-DAVE and Hedvig HRICAK. “Role of MRI in prostate cancer detection”. In: *NMR in Biomedicine* 27.1 (2014), pp. 16–24.

ción sobre la textura, anatomía, vascularización y densidad celular de la próstata, permitiendo caracterizar lesiones en diferentes zonas de la próstata, incluso en lugares donde no han podido ser localizadas mediante él (DRE) e incluso en la biopsia guiada por ecografía transrectal ⁵²¹.

Para el análisis y reporte de hallazgos malignos sobre este tipo de imágenes radiológicas, los expertos utilizan el protocolo *Prostate Imaging Reporting and Data System* (PI-RADS 2.1). Este protocolo de graduación estandariza la interpretación visual de la mp-MRI, informando puntuaciones según la sospecha de malignidad de la lesión de interés observada estimando un valor dependiendo de la malignidad del hallazgo: baja (categoría 1 o 2 de PI-RADS), intermedia (categoría 3 de PI-RADS) y alta (categoría 4 o 5 de PI-RADS) ²². Por otra parte, como alternativa a la mp-MRI, están las secuencias bi-paramétricas (bp-MRI), las cuales consideran únicamente imágenes T2W, DWI y mapas ADC ²³. Esta configuración bp-MRI mejora algunas limitaciones de la mp-MRI como lo son, el tiempo de adquisición, su elevado costo y el uso de agentes de contraste a base de gadolinio que podrían generar efectos secundarios no deseados en los pacientes ²⁴. Además, estudios en barridos de tamizaje poblacional han reportado que los hallazgos clínicos obtenidos utilizando las imágenes de estudios bp-MRI, han sido similares con respecto a utilizar mp-MRI ²⁴. Por consiguiente, bp-MRI es una alternativa versátil que ofrece eficiencia sin perder ventajas en la tarea de caracterización de lesiones csPCa. En este trabajo se optó por utilizar este tipo de secuencias, a continuación se da una breve descripción de los parámetros que componen las secuencias bp-MRI.

²¹ Linda M JOHNSON et al. “Multiparametric MRI in prostate cancer management”. In: *Nature reviews Clinical oncology* 11.6 (2014), pp. 346–353.

²² Thomas BEYER et al. “PI-RADS 2.1–Image Interpretation: The Most Important Updates and Their Clinical Implications”. In: *RöFo-Fortschritte auf dem Gebiet der Röntgenstrahlen und der bildgebenden Verfahren*. Vol. 193. 07. Georg Thieme Verlag KG. 2021, pp. 787–796.

²³ Karen Cecilie Duus THESTRUP et al. “Biparametric versus multiparametric MRI in the diagnosis of prostate cancer”. In: *Acta radiologica open* 5.8 (2016), p. 2058460116663046.

²⁴ Michele SCIALPI et al. “Biparametric MRI of the prostate”. In: *Turkish journal of urology* 43.4 (2017), p. 401.

1.1.1. Imagen ponderada en T2 (T2W). Esta modalidad permite la observación de la anatomía de la glándula prostática ⁵. Para obtener esta secuencia, se aplica un campo magnético que interactúa con los protones de hidrógeno y otros tejidos del cuerpo en estudio. Para esto, inicialmente el cuerpo es sometido a un campo magnético, construyendo una representación de la anatomía molecular en el tejido, permitiendo visualizar la morfología de la glándula prostática ²⁵. Debido a su resolución espacial, resolución de contraste y su visualización multiplanar, se facilita la detección de lesiones en diferentes perspectivas (transaxial, coronal y sagital) ²⁶.

En secuencias T2W, la zona periférica se presenta normalmente hiperintensa, lo cual facilita la identificación de hallazgos de lesión como focos hipointensos. En la figura 1.a) se presenta un ejemplo de lesión csPCa la cual se visualiza como una región hipointensa (oscura) en la zona periférica. En la misma imagen también se señala una lesión sana en la zona periférica, la cual se visualiza como una región hiperintensa (clara). Sin embargo, no siempre distinguir estos patrones es una tarea fácil para los expertos. Por ejemplo, comúnmente las lesiones csPCa se pueden confundir con hallazgos que no tienen una relación con el cáncer, tales como, hemorragia intraglandular, hiperplasia prostática benigna (HPB) o prostatitis ²⁷. En la Figura 1.b) se ilustra uno de estos casos, donde se observa un tejido sano con una intensidad hipointensa ²⁸. Por último, se presenta una próstata sana en la Figura 2.c).

²⁵ R SÁNCHEZ-ORO et al. “Resonancia magnética de próstata: guía práctica de interpretación e informe según PI-RADS versión 2.1”. In: *Radiología* 62.6 (2020), pp. 437–451.

²⁶ Nathan LAWRENTSCHUK et al. “‘Prostatic evasive anterior tumours’: the role of magnetic resonance imaging”. In: *BJU international* 105.9 (2010), pp. 1231–1236.

²⁷ Gary J KELLOFF; Peter CHOYKE, and Donald S COFFEY. “Challenges in clinical prostate cancer: role of imaging”. In: *American journal of roentgenology* 192.6 (2009), pp. 1455–1470.

²⁸ Jorge A OCANTOS; Marcelo Andrés PIETRANI, and Lisandro PAGANINI. “Resonancia magnética de próstata: morfología y metabolismo”. In: *Revista Argentina de Urología* 72.4 (2007), p. 193.

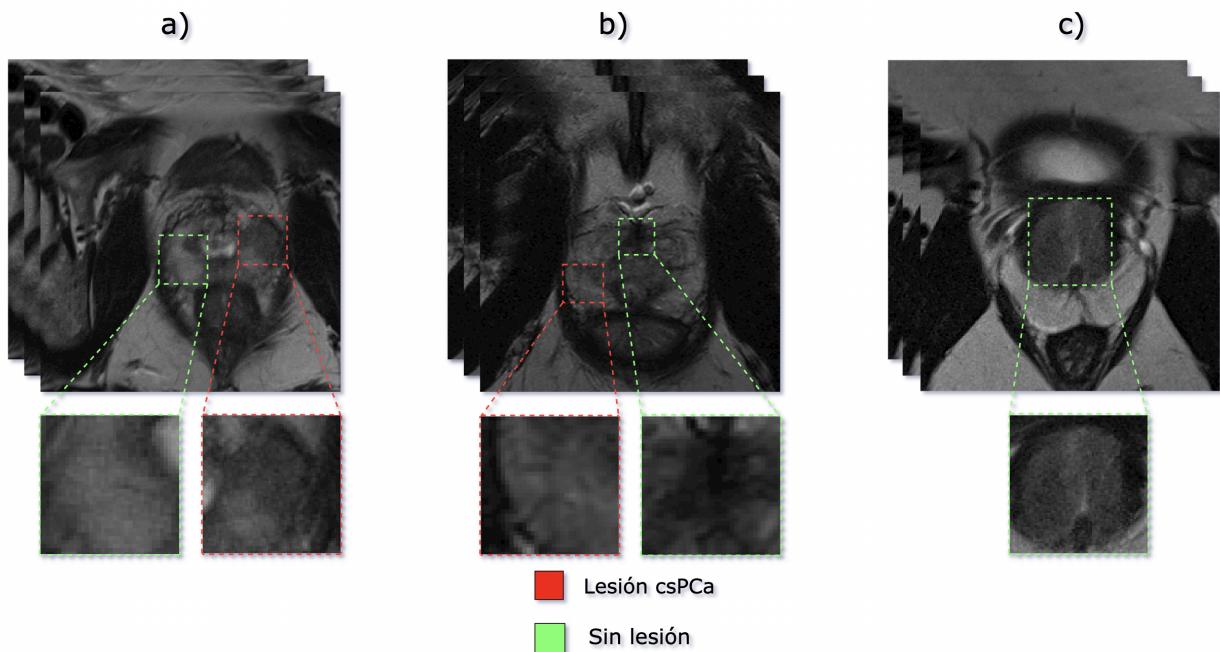


Figura 1

Ejemplos de hallazgos en secuencias T2W. a) Se presenta una lesión maligna y un tejido sano que se observan como regiones hipointensas e hiperintensas en la zona periférica. b) Muestra un tejido sano con una intensidad hipointensa y una lesión csPCa que no se observa hipointensa. c) Se presenta un estudio de una glándula prostática sin lesiones.

1.1.2. Imagen ponderada en difusión (DWI) y mapas ADC. La Imagen de Difusión Ponderada (DWI) aprovecha la propiedad física de la difusión, que describe el movimiento browniano o aleatorio de las moléculas de agua en los tejidos ²⁹. Esta técnica resulta sensible al movimiento de moléculas, especialmente en el espacio extracelular, siendo así de particular interés para el análisis de lesiones relacionadas con el cáncer. Este movimiento aleatorio molecular está acotado principalmente por la densidad celular ³⁰. En regiones de la glándula prostática donde normalmente se encuentran tejidos sanos, las moléculas de agua tienen una difusión libre.

²⁹ Olaf DIETRICH et al. “Technical aspects of MR diffusion imaging of the body”. In: *European journal of radiology* 76.3 (2010), pp. 314–322.

³⁰ Dow-Mu KOH and David J COLLINS. “Diffusion-weighted MRI in the body: applications and challenges in oncology”. In: *American Journal of Roentgenology* 188.6 (2007), pp. 1622–1635.

Por el contrario, en zonas donde el movimiento de las moléculas está limitado o hay difusión restringida, normalmente habitan lesiones csPCa ³¹. De esta forma, en la secuencia DWI, la difusión de las moléculas sigue una relación con el grado de atenuación de la señal, de modo que zonas con una señal hiperintensa se relacionan con una difusión restringida, lo que se traduce a una posible lesión csPCa. Por otra parte, las regiones hipointensas se relacionan comúnmente en los protocolos clínicos con tejidos benignos ³². En la Figura 2.a) se presenta una lesión csPCa con patrones hiperintensos. En la Figura 2.b) se observan dos regiones hiperintensas; sin embargo, solo una de ellas corresponde a una lesión csPCa. En la Figura 2.c) se presenta una glándula prostática totalmente sana.

³¹ Iztok CAGLIC and Tristan BARRETT. “Diffusion-weighted imaging (DWI) in lymph node staging for prostate cancer”. In: *Translational Andrology and Urology* 7.5 (2018), p. 814.

³² Pilar Vallejo DESVIAT et al. “Difusión en resonancia magnética en el estudio de lesiones hepáticas”. In: *Cirugía Española* 91.1 (2013), pp. 9–16.

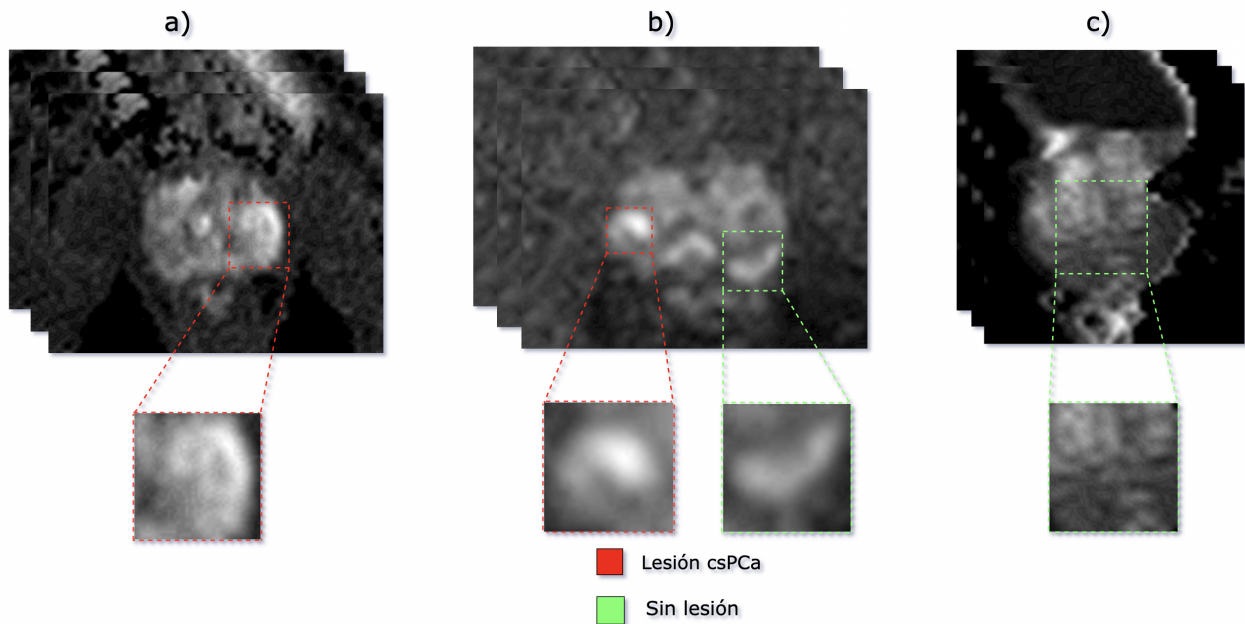


Figura 2

Ejemplos de hallazgos en secuencias DWI. a) Se presenta una lesión maligna que se observa como una región hiperintensa. b) Muestra un tejido sano y otro con lesión csPCa, ambos con características malignas (hiperintenso). c) Se presenta un estudio de una glándula prostática sin lesiones.

En la secuencia DWI existe particularmente el parámetro b que es ajustable en la obtención de la MRI y representa la intensidad con la cual se aplican gradientes en la adquisición del estudio. La modalidad DWI se obtiene con diferentes valores de b , permitiendo detectar posibles anomalías en la glándula por medio de la obtención de los mapas de Coeficiente de Difusión Aparente (ADC)³³. De esta manera, los mapas ADC representan cuantitativamente el desplazamiento de las moléculas de agua, mostrando la difusión de los tejidos de interés³⁴. Por tanto, regiones con difusión restringida se observan como zonas hipointensas así como se ilustra

³³ Roland BAMMER. “Basic principles of diffusion-weighted imaging”. In: *European journal of radiology* 45.3 (2003), pp. 169–184.

³⁴ Hersh CHANDARANA and Bachir TAOULI. “Diffusion and perfusion imaging of the liver”. In: *European journal of radiology* 76.3 (2010), pp. 348–358.

en la Figura 3.a) ³². No obstante, esta secuencia presenta una alta variabilidad, encontrando regiones sanas que se asemejan a hallazgos malignos (ver Figura 3.b). Los mapas ADC han sido esenciales para discriminar entre tejidos malignos o normales. No obstante, una de sus limitaciones radica en su baja resolución en comparación con las otras secuencias, lo que puede dificultar una adecuada caracterización de malignidad de las lesiones para diferentes expertos

35.

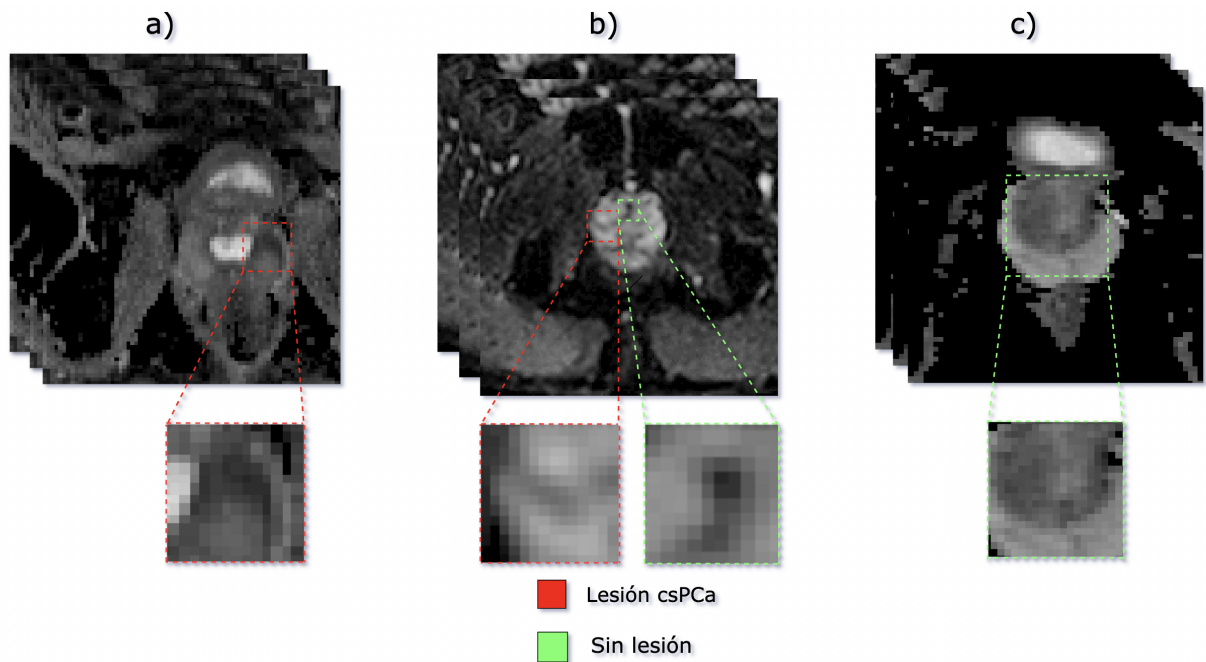


Figura 3

Ejemplos de hallazgos en secuencias ADC. a) Se presenta una lesión maligna que se observa como una región hipointensa. b) Muestra un tejido sano y otro con lesión csPCa, ambos con características malignas (hipointenso). c) Se presenta un estudio de una glándula prostática sin lesiones.

³⁵ Amol Madanlal LAHOTI et al. “Role and application of diffusion-weighted imaging in evaluation of prostate cancer”. In: *Indian Journal of Medical and Paediatric Oncology* 39.03 (2018), pp. 349–354.

1.2. REPRESENTACIONES EMBEBIDAS

El aprendizaje de representaciones se refiere al proceso de ajustar una función paramétrica $f_\theta : \Omega \rightarrow \mathbf{Z}$ que permita transformar un dato de entrada $\mathbf{X} \in \Omega$, en nuestro contexto una imagen o volumen de información MRI, en un vector de características o tensor $\mathbf{z} \leftarrow f_\theta(\mathbf{X}) \in \mathbf{Z}$, el cual mejore o facilite la tarea de discriminación que inicialmente se tenga sobre los datos de entrada ³⁶. Estas representaciones, usualmente también referenciadas como representaciones embebidas, embebidos, o vectores de baja dimensionalidad, han sido codificadas de diferentes formas. Recientemente, los métodos de aprendizaje de máquina han demostrado éxito por su habilidad para aprender y extraer características relevantes para el problema utilizando múltiples capas de aprendizaje profundo ³⁶. Diferentes propuestas para el aprendizaje de representaciones han mostrado su efectividad para capturar las características relevantes de los datos, codificando esta información en un espacio embebido más discriminativo ³⁷³⁸. Además de la extracción de la información, en el estado del arte existen diversas estrategias para analizar y procesar estas representaciones embebidas.

1.2.1. Aprendizaje Contrastivo Supervisado (SLC). Una de las técnicas ampliamente utilizadas en el aprendizaje de representaciones embebidas es el *Aprendizaje Contrastivo*. Esta técnica ha sido implementada para generar un espacio discriminativo de representaciones profundas mediante la utilización de una función de pérdida que busca maximizar la distancia entre embebidos que provienen de datos de diferente clase y minimizar la distancia entre embebidos que provienen de datos que pertenecen a una misma clase (ver figura 4). Por ende, al tener

³⁶ Phuc H LE-KHAC; Graham HEALY, and Alan F SMEATON. “Contrastive representation learning: A framework and review”. In: *Ieee Access* 8 (2020), pp. 193907–193934.

³⁷ Qi LIU; Matt J KUSNER, and Phil BLUNSOM. “A survey on contextual embeddings”. In: *arXiv preprint arXiv:2003.07278* (2020).

³⁸ Haochen CHEN et al. “A tutorial on network embeddings”. In: *arXiv preprint arXiv:1808.02590* (2018).

embebidos que se diferencian entre sí, se ha mostrado cómo esta técnica facilita finalmente la tarea de clasificación sobre los datos de entrada ³⁹⁴⁰.

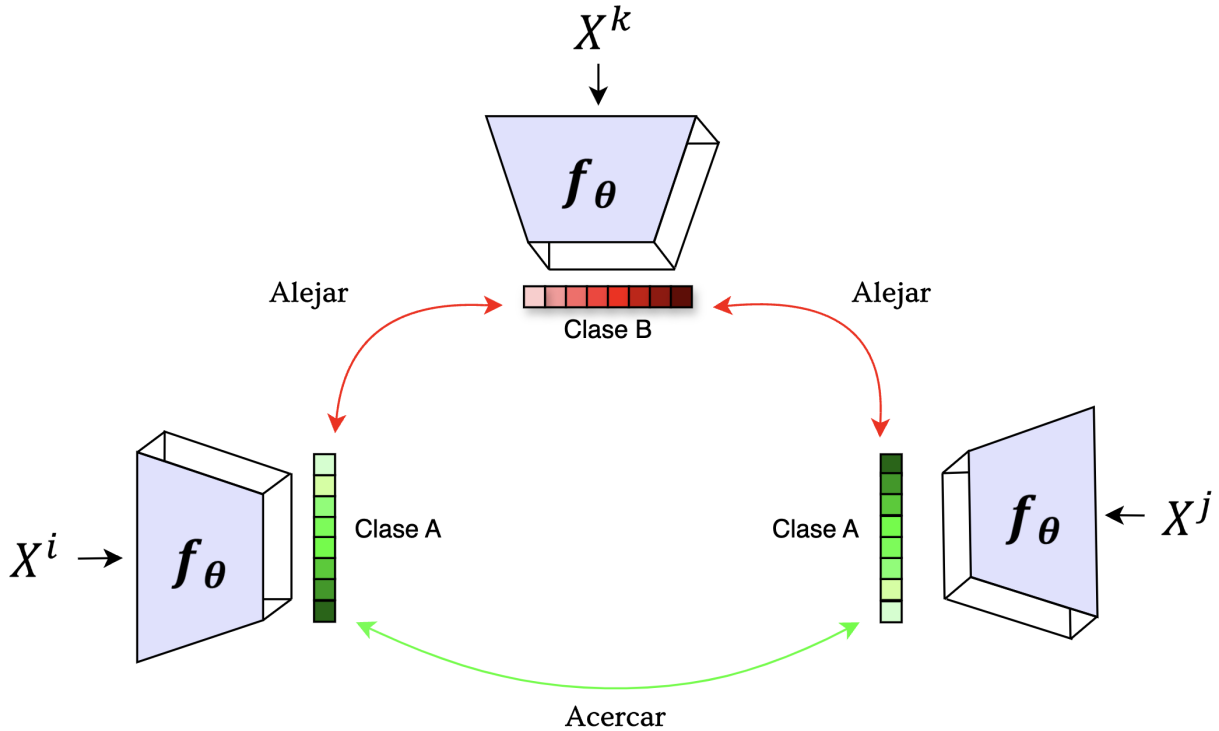


Figura 4

Esquema de aprendizaje contrastivo. Se presenta un espacio embebido contrastivo donde X^i y X^j son muestras de la clase A y X^k una muestra de la clase B. A partir de estas representaciones, f_θ se ajusta para codificar un espacio latente discriminativo mediante la minimización de embebidos proyectados de muestras de la misma clase y la maximización de distancia entre embebidos proyectados de muestras de clase opuesta.

A partir de la formulación del aprendizaje contrastivo, se han propuesto variaciones en su implementación dependiendo del contexto del problema. El esquema que se ilustra en

³⁹ Yesid GUTIÉRREZ; John AREVALO, and Fabio MARTÍNEZ. “Multimodal Contrastive Supervised Learning to Classify Clinical Significance MRI Regions on Prostate Cancer”. In: *2022 44th Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC)*. IEEE. 2022, pp. 1682–1685.

⁴⁰ Yesid GUTIÉRREZ; John AREVALO, and Fabio MARTÍNEZ. “A contrastive weakly supervised learning to characterize malignant prostate lesions in BP-MRI”. in: *Biomedical Signal Processing and Control* 96 (2024), p. 106584.

la Figura 4 corresponde a un enfoque contrastivo supervisado. En esta ilustración, dado un conjunto de datos (*Batch*, término común en inglés) \mathcal{B} con N muestras y labels $(X_i, Y_i)_{i=1}^N$, se propone inicialmente un aumento de datos que permita crear múltiples vistas de una misma muestra. Estas usualmente corresponden a rotaciones, zooms, recortes aleatorios, entre otros. De esta manera se obtiene un conjunto aumentado $\mathcal{B}_{Aug} = (\tilde{X}_i, \tilde{Y}_i)_{i=1}^{2N}$, donde $\tilde{X} = Aug(X)$. Acá \tilde{X}_{2N-1} y \tilde{X}_{2N} son versiones aumentadas de X_N . Posteriormente utilizando una red de encoder y una red de proyección, cada muestra es ahora proyectada en un vector embebido $\mathbf{x} = Proj(Enc(X))$, donde $Enc(\cdot)$ usualmente corresponde a un extractor de características profundas y $Proj(\cdot)$ a una red con múltiples capas densas ⁴¹. Para su ajuste, este enfoque utiliza una función de costo que busca minimizar la distancia entre pares de muestras de igual clase $\mathbf{x}_i, \mathbf{x}_p$ (par ancla-positivo) y maximizar la distancia con respecto a muestras adversarias $\mathbf{x}_i, \mathbf{x}_a$ (par ancla-negativo), es decir, de diferente clase. En este orden de ideas, la función de pérdida sigue de la siguiente regla:

$$\mathcal{L}_{NT-Xent} = -\frac{1}{|I|} \sum_{i \in I} \frac{1}{|P(i)|} \sum_{p \in P(i)} \log \frac{\exp(sim(\mathbf{x}_i, \mathbf{x}_p)/\tau)}{\sum_{a \in I} \exp(sim(\mathbf{x}_i, \mathbf{x}_a)/\tau)} \quad (1)$$

Donde I es el conjunto de índices de \mathcal{B}_{Aug} , es decir, $I = \{1, 2, \dots, 2N\}$. Además, $P(i) = \{p \in I \setminus \{i\} : \tilde{Y}_p = \tilde{Y}_i\}$, es decir, consiste en un conjunto de muestras "positivas", esto es, índices de muestras de la misma clase que el ancla \mathbf{x}_i . Además cada a , son los índices donde su clase es diferente al ancla $a \in I : \tilde{Y}_a \neq \tilde{Y}_i$. Por otra parte, $sim(\cdot)$ corresponde a una medida de similaridad entre vectores embebidos y τ representa un hiper-parámetro de temperatura que ajusta la distribución de similitudes entre los vectores de características ⁴². La principal noción de esta función de pérdida es obtener una óptima representación embebida discriminativa,

⁴¹ Ting CHEN et al. "A simple framework for contrastive learning of visual representations". In: *International conference on machine learning*. PMLR. 2020, pp. 1597–1607.

⁴² Prannay KHOSLA et al. "Supervised contrastive learning". In: *Advances in neural information processing systems* 33 (2020), pp. 18661–18673.

agrupando representaciones de elementos positivos, a partir de un conjunto de elementos que contiene muestras de la clase opuesta. Como resultado se obtiene un espacio de representaciones embebidas robusto, que en escenarios con carencia de datos etiquetados ha demostrado tener un rendimiento sobresaliente³⁹⁴¹. El enfoque contrastivo ha demostrado además, ser una estrategia versátil en diversos contextos, por ejemplo en la clasificación de imágenes comunes, imágenes médicas e incluso en tareas de segmentación⁴³⁴⁴⁴⁵.

1.2.2. Representaciones geométricas SPD. En diferentes aplicaciones de visión por computador se ha evidenciado la efectividad de las matrices simétricas definidas positivas (SPD) como un conjunto de representaciones alternativo para la compactación de información relevante a partir de un banco de características. Estos embebidos pueden generalizar el aprendizaje de patrones relevantes en la tarea de interés, incluso en ambientes con datos limitados⁴⁶. Específicamente, Dado un banco de características $\mathbf{F} = \{F^{(i)}\}_{i=1}^N$, cada una de ellas siendo un vector, matriz o tensor, estas se pueden resumir en una matriz simétrica definida positiva (SPD) \mathbf{M} de dimensión $N \times N$. Específicamente, codificando una matriz \mathbf{C} cuyas filas sean dichas características $F^{(i)}$ vectorizadas y computando $\mathbf{M} = \mathbf{C}\mathbf{C}^\top$, se obtiene una matriz que resulta simétrica, pues $\mathbf{M} = \mathbf{M}^\top$ y también se puede ver que es definida positiva. El conjunto de todas estas matrices, denotado como $Sym_{++}^N = \{\mathbf{M} \in \mathbb{R}^{N \times N} \mid \mathbf{M} = \mathbf{M}^\top \wedge \forall v \neq 0, v\mathbf{M}v^\top > 0\}$ forma un espacio geométrico, una variedad diferenciable dada por las matrices simétricas definidas

⁴³ Peng WANG et al. “Contrastive learning based hybrid networks for long-tailed image classification”. In: *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*. 2021, pp. 943–952.

⁴⁴ Yuhao ZHANG et al. “Contrastive learning of medical visual representations from paired images and text”. In: *Machine Learning for Healthcare Conference*. PMLR. 2022, pp. 2–25.

⁴⁵ Yawen WU et al. “Distributed contrastive learning for medical image segmentation”. In: *Medical Image Analysis* 81 (2022), p. 102564.

⁴⁶ Yesid GUTIÉRREZ et al. “A Multimodal Geometric Deep Representation to Support Bi-Parametric Prostate Cancer Lesion Classification”. In: *2023 IEEE 20th International Symposium on Biomedical Imaging (ISBI)*. IEEE. 2023, pp. 1–4.

positivas. Específicamente, esta es una variedad Riemanniana, curva, no Euclidiana, con una métrica Riemanniana que brinda una distancia entre elementos en este espacio ⁴⁷. Al considerar estas particularidades geométricas sobre Sym_{++}^N , diferentes trabajos han mostrado que aplicar un procesamiento que preserve la geometría de los descriptores SPD, permite alcanzar mejores resultados en diferentes tareas de clasificación y reconocimiento de patrones ⁴⁸. Actualmente, en la literatura existen diferentes herramientas de aprendizaje en este espacio embebido donde se aprovecha la geometría Riemanniana, formulando capas de aprendizaje profundo que contemplen y respeten las características geométricas de los embebidos SPD ⁴⁹⁵⁰.

1.3. SISTEMAS DE DIAGNÓSTICO ASISTIDO POR COMPUTADOR

En el estado del arte, las estrategias de aprendizaje automático han mejorado la precisión en el diagnóstico y clasificación de lesiones csPCa ⁵¹. Algunos de estos métodos, han propuesto estrategias de aprendizaje supervisadas para la clasificación de patrones malignos a partir de la extracción de características profundas de las imágenes mediante la implementación de redes profundas convolucionales. Por ejemplo, Mehrtash *et.al* propusieron un enfoque multimodal que a partir de parches volumétricos procesa diferentes secuencias, como mapas de ADC, BVAL y la secuencia de contraste K^{trans} , utilizando convoluciones 3D e incluyendo un módulo para la

⁴⁷ William M BOOTHBY. *An introduction to differentiable manifolds and Riemannian geometry*. Academic press, 1986.

⁴⁸ Oncel TUZEL; Fatih PORIKLI, and Peter MEER. “Pedestrian detection via classification on riemannian manifolds”. In: *IEEE transactions on pattern analysis and machine intelligence* 30.10 (2008), pp. 1713–1727.

⁴⁹ Vincent ARSIGNY et al. “Geometric means in a novel vector space structure on symmetric positive-definite matrices”. In: *SIAM journal on matrix analysis and applications* 29.1 (2007), pp. 328–347.

⁵⁰ Zhiwu HUANG and Luc VAN GOOL. “A riemannian network for spd matrix learning”. In: *Thirty-first AAAI conference on artificial intelligence*. 2017.

⁵¹ Islam REDA et al. “Deep learning role in early diagnosis of prostate cancer”. In: *Technology in cancer research & treatment* 17 (2018), p. 1533034618775530.

integración de información zonal de la región analizada ⁵². Otros autores, también han propuesto el uso de estrategias de aprendizaje por transferencia (*transfer learning*) para soportar la clasificación de lesiones clínicamente significativas en contextos con limitación de datos ⁵³. En general, la integración de múltiples secuencias MRI a través de arquitecturas convolucionales ha demostrado mejoras en clasificación y la detección de lesiones csPCa ^{54,52}. No obstante, estos métodos siguen dependiendo de la disponibilidad de grandes volúmenes de datos de MRI etiquetados, considerando además que los datos deben ser preferiblemente balanceados. Lo mencionado constituye un problema dada la alta variabilidad en datos clínicos reales y la limitación de datos anotados ⁵⁵. Para esto, trabajos anteriores han propuesto utilizar redes preentrenadas ⁵⁶. Sin embargo, estas redes contienen un alto número de parámetros, el cual en enfoques con pocos datos puede llevar a un sobreajuste. Por otra parte, Weißer *et.al* propusieron una estrategia de aprendizaje débilmente supervisada soportada por anotaciones radiológicas a nivel de cortes ⁵⁷. Sin embargo, se evidenció la necesidad de usar más datos para un correcto ajuste del modelo.

⁵² Alireza MEHRTASH et al. “Classification of clinical significance of MRI prostate findings using 3D convolutional neural networks”. In: *Medical imaging 2017: computer-aided diagnosis*. Vol. 10134. SPIE. 2017, pp. 589–592.

⁵³ Quan CHEN et al. “A transfer learning approach for classification of clinical significant prostate cancers from mpMRI scans”. In: *Medical imaging 2017: Computer-aided diagnosis*. Vol. 10134. SPIE. 2017, pp. 1154–1157.

⁵⁴ Nader ALDOJ et al. “Semi-automatic classification of prostate cancer on multi-parametric MR imaging using a multi-channel 3D convolutional neural network”. In: *European radiology* 30.2 (2020), pp. 1243–1253.

⁵⁵ S LIU et al. “Prostate Cancer Diagnosis using Deep Learning with 3D Multiparametric MRI. SPIE Med”. In: *Imaging* 10134 (2017), pp. 1–4.

⁵⁶ William Olurotimi FALANA; Ali SERENER, and Sertan SERTE. “Clinically Significant Prostate Cancer Classification Using Anatomical T2W MRI Sequences and Deep Learning”. In: *2023 7th International Symposium on Multidisciplinary Studies and Innovative Technologies (ISMSIT)*. IEEE. 2023, pp. 1–6.

⁵⁷ Cedric WEISSER et al. “Weakly supervised MRI slice-level deep learning classification of prostate cancer approximates full voxel-and slice-level annotation: Effect of increasing training set size”. In: *Journal of Magnetic Resonance Imaging* 59.4 (2024), pp. 1409–1422.

En la literatura, recientemente los trabajos se han enfocado en métodos que buscan mejorar la detección de lesiones csPCa. Por ejemplo, Saha *et.al* implementó mecanismos de atención y un clasificador residual para soportar la reducción de falsos positivos ⁵⁸. Así mismo, en la literatura se pruso la NRD-Net, una red de destilación robusta al ruido para la detección de lesiones csPCa en secuencias bp-MRI. Este esquema propone un módulo de segmentación y un modelo de clasificación de múltiples ramas ⁵⁹. Otros enfoques han propuesto el cálculo de características radiómicas a partir de los estudios bp-MRI para mejorar la detección de metástasis de cáncer de próstata ⁶⁰. También, para mejorar la detección, Coen de Vente *et.al* propusieron la incorporación de mapas de segmentación en distintos niveles de la arquitectura, además de concatenar segmentaciones de diferentes zonas de la próstata como información previa, lo que demostró mejoras en la sensibilidad para la detección de lesiones csPCa ⁶¹. A pesar del buen desempeño de estas redes, es ampliamente conocido su limitación en escenarios con pocos datos, optando por enfoques 2D para tener mas ejemplos durante el entrenamiento. Sin embargo, de esta manera se ignora la información 3D que ha demostrado contribuir de mejor forma a la caracterización de las lesiones ⁶². Por otra parte, se ha evidenciado cómo el rendimiento de estas redes decae al ser validadas en nuevos conjuntos de datos, lo que demuestra una falta de

-
- ⁵⁸ Anindo SAHA; Matin HOSSEINZADEH, and Henkjan HUISMAN. “End-to-end prostate cancer detection in bpMRI via 3D CNNs: effects of attention mechanisms, clinical priori and decoupled false positive reduction”. In: *Medical image analysis* 73 (2021), p. 102155.
- ⁵⁹ Xiangtong DU et al. “NRD-Net: a noise-resistant distillation network for accurate diagnosis of prostate cancer with bi-parametric MRI images”. In: *Multimedia Tools and Applications* 83.11 (2024), pp. 33597–33614.
- ⁶⁰ Song XINYANG et al. “A machine learning radiomics model based on bpMRI to predict bone metastasis in newly diagnosed prostate cancer patients.” In: *Magnetic Resonance Imaging* 107 (2024), pp. 15–23.
- ⁶¹ Coen DE VENDE et al. “Deep learning regression for prostate cancer detection and grading in bi-parametric MRI”. in: *IEEE Transactions on Biomedical Engineering* 68.2 (2020), pp. 374–383.
- ⁶² Peter REMEIJER et al. “3-D portal image analysis in clinical practice: an evaluation of 2-D and 3-D analysis techniques as applied to 30 prostate cancer patients”. In: *International Journal of Radiation Oncology* Biology* Physics* 46.5 (2000), pp. 1281–1290.

generalización, indicando un sobre ajuste de representaciones a una misma distribución limitada de datos.

Recientemente, estrategias de aprendizaje profundo han sido propuestas buscando robustez en escenarios con carencia de datos. Un ejemplo es el método propuesto por Gutierrez *et.al*, el cual presenta un marco de aprendizaje contrastivo supervisado (SCL) para soportar el aprendizaje de un espacio de representaciones robusto ³⁹. A pesar de su buen rendimiento, estas representaciones aprendidas ignoran la variabilidad persistente en las secuencias MRI. Para esto, en la literatura se ha propuesto descriptores compactos SPD para la codificación de características extraídas de las secuencias bp-MRI, que posteriormente son utilizados en la clasificación de lesiones csPCa ⁴⁶. Particularmente, este enfoque demostró un rendimiento sobresaliente al entrenar con pocos datos; sin embargo, su método se enfoca en ajustar una pérdida guiada por las etiquetas, en lugar de considerar un aprendizaje enfocado en el espacio embebido de dichas representaciones.

2. PROBLEMA DE INVESTIGACIÓN

El cáncer de próstata (PCa) es el cuarto tipo de cáncer mas diagnosticado a nivel mundial, y cuenta con mas de 1.4 millones de nuevos casos en el 2020 ². En Colombia el PCa es la segunda causa principal de muertes por cáncer en hombres ⁶³. El diagnóstico del PCa es apoyado por métodos como el análisis del antígeno prostático (PSA, por sus siglas en inglés) y el examen digital rectal (DRE, por sus siglas en inglés). Sin embargo, estas metodologías reportan una baja especificidad, con un número marcado de falsos positivos⁴. Recientemente el análisis de imágenes diagnósticas mp-MRI/bp-MRI ha permitido una potencial alternativa no-invasiva para caracterizar lesiones relacionadas con el PCa. Sin embargo, el análisis observacional de estas secuencias es una labor subjetiva y compleja, incluso para expertos, debido a la variabilidad en apariencia que hay en lesiones de cáncer de próstata ⁹. Diferentes estrategias computacionales se han propuesto para mitigar la subjetividad y variabilidad diagnóstica. Sin embargo, estas arquitecturas requieren grandes cantidades de datos para obtener un desempeño favorable de clasificación. Además, estas arquitecturas ignoran potenciales descriptores compactos que capturen relaciones de textura, y que puedan constituir un descriptor más discriminativo de las lesiones, ayudando a una mejor caracterización de la malignidad.

Pregunta de Investigación

¿Cómo implementar un enfoque de aprendizaje profundo para la clasificación de lesiones clínicamente significativas de cáncer de próstata (csPCa) a partir de una caracterización geométrica de embebidos sobre estudios de imágenes de resonancia magnética biparamétrica (bp-MRI)?

⁶³ Hyuna SUNG et al. “Global cancer statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries”. In: *CA: a cancer journal for clinicians* 71.3 (2021), pp. 209–249.

3. OBJETIVOS

Objetivo General

- Implementar un modelo de aprendizaje profundo para la clasificación de vectores embebidos que representan lesiones de próstata desde observaciones MRI

Objetivos Específicos

- Adaptar un conjunto de estudios bp-MRI que incluya anotaciones asociadas a lesiones sanas y clínicamente significativas.
- Diseñar una estrategia que codifique vectores embebidos a partir de los hallazgos en secuencias bp-MRI.
- Desarrollar una arquitectura de aprendizaje profundo que incluya la representación de vectores embebidos para la clasificación de lesiones clínicamente significativas.
- Evaluar la estrategia desarrollada en cuanto a la capacidad de clasificar lesiones clínicamente significativas de cáncer de próstata.

4. MÉTODO PROPUESTO

En este trabajo se implementó una estrategia de aprendizaje contrastivo supervisado a partir de información multimodal bp-MRI que permitió ajustar, explorar y validar un espacio embebido de baja dimensionalidad para clasificar lesiones csPCa. Este método inicialmente implementó una arquitectura 3D convolucional para la extracción de información multimodal imegeneológica a través de ramas independientes por modalidad, que posteriormente son fusionadas para ser compactadas en vectores embebidos, que buscan crear un espacio de representaciones que se ajuste de acuerdo a la malignidad. Estas representaciones son proyectadas a clasificadores para obtener una predicción de clase. En la Figura 5 se presenta un esquema del método propuesto.

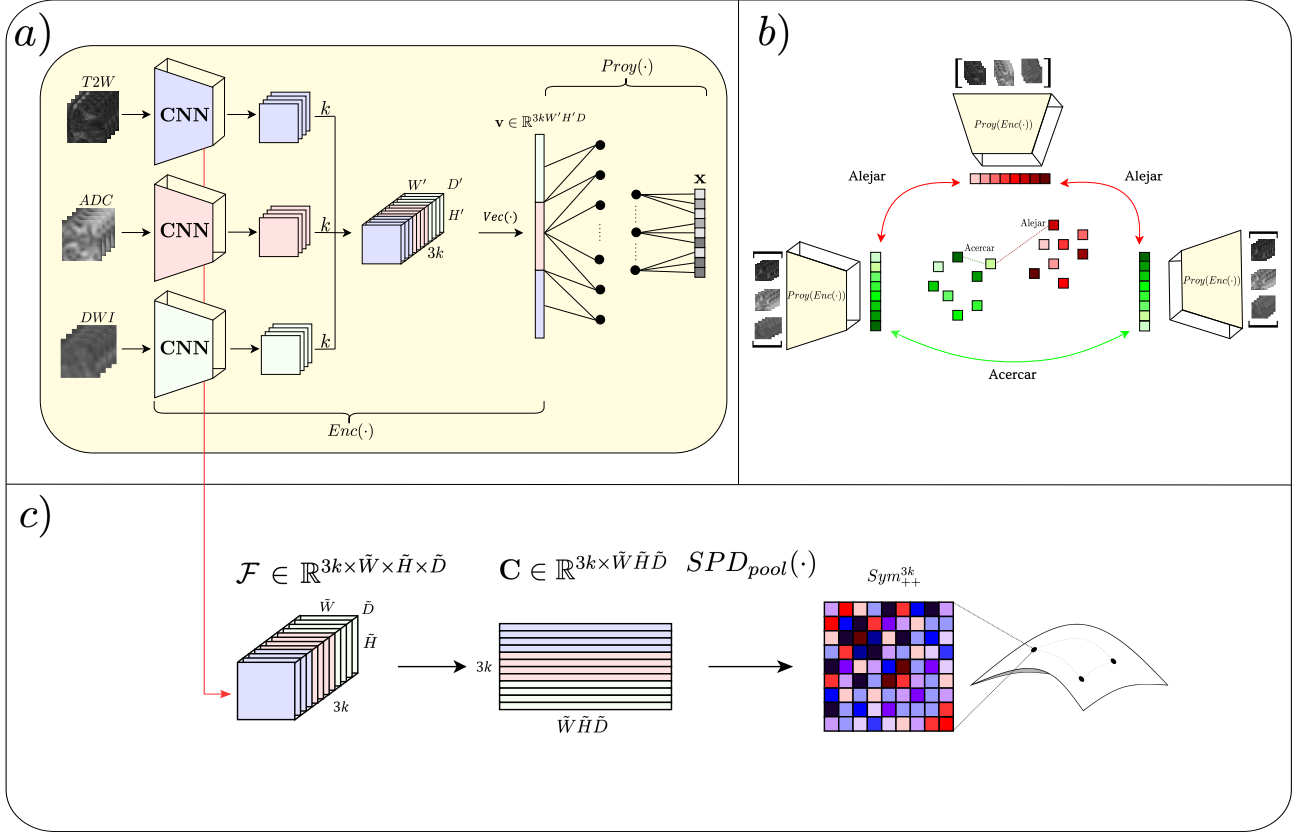


Figura 5

Pipeline del método propuesto. En **a)** se presenta la arquitectura $Enc(\cdot)$ y $Proj(\cdot)$ la cual, recibe las tres modalidades de MRI para codificar un vector embebido \mathbf{x} . **b)** Posteriormente, la arquitectura se rige por un enfoque contrastivo supervisado. **c)** En una segunda etapa, se codifican matrices SPD las cuales se procesan y clasifican.

4.1. ESQUEMA CONVOLUCIONAL 3D PARA LA EXTRACCIÓN DE CARACTERÍSTICAS MRI

Se consideró una arquitectura multimodal bp-MRI que recibe tres parches volumétricos por parámetro, considerándolos como: \mathbf{X}_{T2} , \mathbf{X}_{ADC} , \mathbf{X}_{HBV} . Cada parche tiene una dimensión de $W \times H \times D$, donde $W \times H$ es la dimensión espacial y D es el número de cortes disponible en cada estudio bp-MRI. En este orden de ideas, cada parche bp-MRI se incorpora a una estructura convolucional que contiene tres ramas independientes Enc^{T2} , Enc^{ADC} , Enc^{HBV} ,

cada una de estas corresponde a una misma arquitectura, la cual en este trabajo representa una red convolucional 3D para aprovechar la información espacial y morfológica de las lesiones (ver figura 5.a). Cada una de estas ramas retorna un banco con k características profundas por modalidad, esto es $\mathbf{F}_{T2}, \mathbf{F}_{ADC}, \mathbf{F}_{HBV} \in \mathbb{R}^{k \times W' \times H' \times D'}$, donde $W' \times H' \times D'$ son las dimensiones del volumen resultante y k es el número de activaciones correspondientes a los volúmenes resultantes. Posteriormente, estos bancos de características profundas son concatenados en un único banco $\mathbf{F} \in \mathbb{R}^{3k \times W' \times H' \times D'}$ para posteriormente ser vectorizado $\mathbf{v} = \text{vec}(\mathbf{F}) \in \mathbb{R}^{3kW'H'D'}$. Así, esta representación constituye un único encoder $Enc(\cdot)$ tal que $\mathbf{v} = \text{vec}(Enc(\mathbf{X}_{T2}, \mathbf{X}_{ADC}, \mathbf{X}_{HBV}))$. Finalmente mediante una red de proyección $Proy(\cdot)$ se calcula un vector embebido que describe el parche multimodal bp-MRI de entrada como una representación de baja dimensión. Este embebido es normalizado como $\mathbf{x} = \frac{Proy(\mathbf{v})}{\|Proy(\mathbf{v})\|}$ para una mejor eficiencia en el entrenamiento del modelo y facilidad en el cálculo de similitudes, formando así una hiper-esfera unitaria⁴¹. El encoder extractor de características y el módulo de proyección $Enc(\cdot)$, $Proy(\cdot)$ respectivamente, son entrenados para la codificación automática de representaciones embebidas de parches bp-MRI con el objetivo de ajustar un espacio latente que facilite la tarea de diferenciar entre vectores que proyectan regiones malignas y vectores que proyectan regiones sanas.

4.2. ESQUEMA DE APRENDIZAJE CONTRASTIVO SUPERVISADO (SCL)

Para la construcción de embebidos que representan lesiones prostáticas se planteó un esquema de aprendizaje contrastivo supervisado, que involucra una selección de la tripleta contrastiva involucrada en la función de pérdida. En el enfoque contrastivo convencional, a partir de un lote de datos aumentado \mathcal{B}_{aug} , cada embebido \mathbf{x} puede tomar el rol de ancla (n anclas), además por cada ancla, se tienen (n) pares positivos y (n) muestras adversarias. De esta manera, la combinación de tripletas contrastivas durante el entrenamiento resulta en un alto costo computacional ($O(n^3)$). Además, se ha mostrado que el cómputo de todas estas tripletas

no necesariamente ayuda a una mejor convergencia del modelo ⁶⁴. Esta redundancia de cálculos de similitudes entre pares de muestras, puede llevar a un sobre ajuste y estropear la codificación de representaciones embebidas. En este orden de ideas, en este trabajo se propuso una estrategia para la selección de pares positivos y negativos, comúnmente conocida como minado de datos. Específicamente, se seleccionó un par positivo y un par negativo (\mathbf{x}_p y \mathbf{x}_a , respectivamente) con el objetivo de incrementar la similaridad entre el ancla y el positivo ($\mathbf{x}_i, \mathbf{x}_p$), mientras se maximiza la variabilidad entre el ancla y el negativo ($\mathbf{x}_i, \mathbf{x}_a$). Dentro de las diversas maneras de escoger un par positivo, se usó la estrategia de minado positivo-fácil que consiste en elegir un \mathbf{x}_p cercano al ancla \mathbf{x}_i y que comparta la misma clase ($\tilde{Y}_p = \tilde{Y}_i$), esto es:

$$\mathbf{x}_p = \arg \min_{\mathbf{x}} d(\mathbf{x}_i, \mathbf{x}) \quad \text{s.t.} \quad \tilde{Y} = \tilde{Y}_i, \quad (2)$$

donde $d(\cdot)$ corresponde a la distancia euclidiana $d(\mathbf{x}_i, \mathbf{x}) = \|\mathbf{x}_i - \mathbf{x}\|_2$. Complementariamente, se escogió otra estrategia de minado de datos para el par negativo denominada negativo-semidifícil donde

$$\mathbf{x}_a = \arg \min_{\mathbf{x}} d(\mathbf{x}_i, \mathbf{x}) \quad \text{s.t.} \quad \tilde{Y} \neq \tilde{Y}_i \quad \text{and} \quad d(\mathbf{x}_i, \mathbf{x}) > d(\mathbf{x}_i, \mathbf{x}_p), \quad (3)$$

la cual sugiere escoger la muestra negativa cercana al ancla pero que su distancia sea mayor a la distancia del ancla al par positivo escogido. Este proceso de minado se ilustra en la Figura 6.

⁶⁴ Hong XUAN; Abby STYLIANOU, and Robert PLESS. “Improved embeddings with easy positive triplet mining”. In: *Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision*. 2020, pp. 2474–2482.

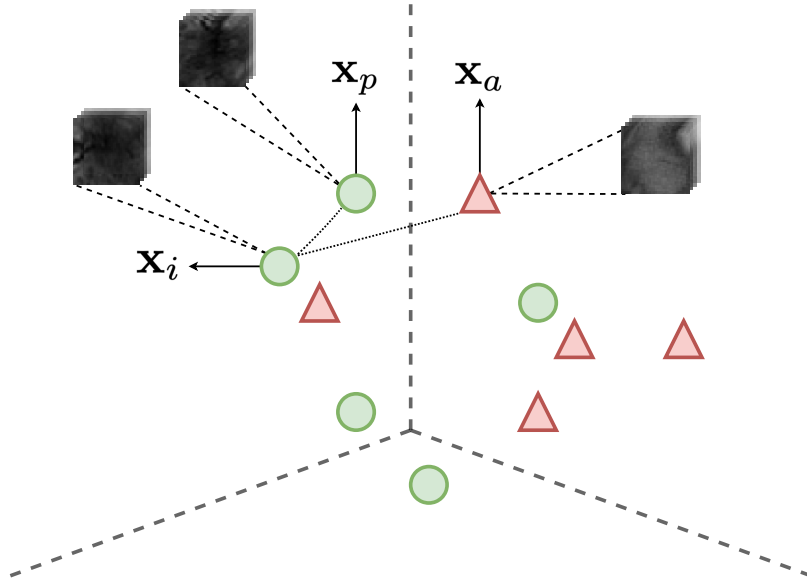


Figura 6

Minado de datos. A partir de un ancla \mathbf{x}_i , la selección del par positivo \mathbf{x}_p y negativo \mathbf{x}_a se sigue mediante una estrategia de minado de datos denominada positivo-fácil y negativo-semidifícil, en la cual se elige la muestra de la misma clase mas cercana al ancla y la muestra de clase contraria mas cercana al ancla cuya distancia sea mayor a la del par positivo escogido respectivamente.

Esta estrategia de minado ha demostrado ser eficaz para filtrar la selección de tuplas suficientemente representativas para el esquema de aprendizaje contrastivo planteado ⁶⁴. Luego de esta selección de la tripleta contrastiva, se sigue un esquema de aprendizaje contrastivo supervisado similar al presentado en la Ecuación (1), en el que se busca minimizar la distancia entre embebidos de igual clase mientras restringe la cercanía de representaciones de diferente clase siguiendo una función de pérdida denominada NT-Xent (*Normalized Temperature-Scaled Cross-Entropy*) Loss ⁴¹, la cual se define como:

$$\mathcal{L}_{NT-Xent} = -\frac{1}{|I|} \sum_{i \in I} \log \frac{\exp(\text{sim}(\mathbf{x}_i, \mathbf{x}_p)/\tau)}{\exp(\text{sim}(\mathbf{x}_i, \mathbf{x}_a)/\tau)}. \quad (4)$$

Donde I es el conjunto de índices de \mathcal{B}_{aug} , \mathbf{x}_i es el ancla seleccionado, \mathbf{x}_p es el par positivo escogido (Ecuación (2)) y \mathbf{x}_a es el par negativo seleccionado en la ecuación 3. Esta pérdida permite ajustar las representaciones embebidas de los parches bp-MRI con el objetivo

de alejar de una manera mas eficiente las proyecciones de parches malignos de parches que provienen de regiones sanas (ver figura 5.b). Existen diversas funciones que han sido propuestas para el aprendizaje contrastivo supervisado, entre ellas, la función de pérdida *Triplet* ha sido una de las primeras propuestas y a su vez ha demostrado éxito en varios escenarios. Es por esto que en este trabajo adicionalmente decidimos evaluar el método considerando esta perdida clásica. Específicamente, esta se define como:

$$\mathcal{L}_{Triplet} = \frac{1}{|I|} \sum_{i \in I} [\|\mathbf{x}_i - \mathbf{x}_p\|_2 - \|\mathbf{x}_i - \mathbf{x}_a\|_2 + \alpha]_+ \quad (5)$$

donde $[\cdot]_+$ es equivalente a la función $\max(0, \cdot)$, $\|\mathbf{x}_i - \mathbf{x}_p\|_2$ representa la distancia euclidiana entre el ancla y el par positivo y $\|\mathbf{x}_i - \mathbf{x}_a\|_2$ representa la distancia entre el ancla y el par negativo. Además, α es un hiper-parámetro conocido como margen y representa la diferencia deseada entre la distancia ancla-positiva y la distancia ancla-negativa⁶⁵⁶⁶.

4.3. CLASIFICACIÓN DE EMBEBIDOS (*CONTRASTIVO*).

A partir del espacio de representaciones aprendido, en una segunda fase, posterior al entrenamiento contrastivo, los vectores embebidos \mathbf{x} fueron evaluados en su capacidad para discriminar lesiones csPCa. Para esto, se ajustaron distintos clasificadores de aprendizaje de máquina sobre el espacio embebido. Particularmente, en este trabajo se utilizaron dos clasificadores: una regresión logística y una máquina de soporte vectorial. A continuación se aborda a detalle cada uno de ellos.

- **Regresión logística:** Este método propone un enfoque estadístico principalmente para problemas de clasificación binaria, permitiendo estimar la probabilidad de dado un vector

⁶⁵ Alexander HERMANS; Lucas BEYER, and Bastian LEIBE. “In defense of the triplet loss for person re-identification”. In: *arXiv preprint arXiv:1703.07737* (2017).

⁶⁶ Weifeng GE. “Deep metric learning with hierarchical triplet loss”. In: *Proceedings of the European conference on computer vision (ECCV)*. 2018, pp. 269–285.

embebido este corresponda a la clase positiva, en este caso, que sea csPCa. A diferencia de otros modelos de regresión que se enfocan en predecir valores continuos, la regresión logística se centra en predecir la probabilidad de pertenencia a una de dos clases posibles⁶⁷⁶⁸. Para este propósito, se utiliza la entropía cruzada binaria, que compara las etiquetas reales con las probabilidades predichas por el modelo. La Regresión Logística esta resulta de modelar $\ln\left(\frac{P}{1-P}\right) = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_k X_k$, donde P es la probabilidad de que ocurra el evento de interés, β_0 es el la intersección o sesgo y $\beta_1, \beta_2, \dots, \beta_k$ son los coeficientes de las variables independientes X_1, X_2, \dots, X_k respectivamente. En el problema de clasificación de lesiones csPCa mediante la regresión logística, estas variables independientes corresponden a los valores en las entradas de cada embebido o vector calculado desde el esquema contrastivo. En este método se utilizan las respectivas etiquetas asociadas a los vectores para ajustar el problema de regresión y posteriormente evaluar la capacidad de las representaciones embebidas de agruparse en un espacio embebido n-dimensional que permiten diferenciar si un caso es una lesión csPCa.

- **Máquina de soporte vectorial:** Es un algoritmo supervisado comúnmente utilizado en problemas de clasificación y regresión, cuyo objetivo es hallar un hiper-plano que maximice el margen entre las clases en un espacio embebido, lo que se logra mediante la resolución de un problema de optimización para datos linealmente separables. Para datos que no son linealmente separables, este algoritmo implementa funciones de kernel, como el kernel gaussiano o polinómico, que permiten mapear los datos a un espacio de mayor dimensión donde la separación lineal sea mas fácil de abordar⁶⁹.

⁶⁷ Michael P LAVALLEY. “Logistic regression”. In: *Circulation* 117.18 (2008), pp. 2395–2399.

⁶⁸ David G KLEINBAUM et al. *Logistic regression*. Springer, 2002.

⁶⁹ William S NOBLE. “What is a support vector machine?” In: *Nature biotechnology* 24.12 (2006), pp. 1565–1567.

4.4. EMBEBIDOS SPD GEOMÉTRICOS

Como alternativa para la clasificación de embebidos extraídos a nivel del enfoque contrastivo, se decidió analizar las características profundas bp-MRI aprendidas en la regla contrastiva del encoder convolucional fueron compactadas en embebidos geométricos para una mayor discriminación entre clases. Dado el esquema propuesto, las características profundas brindan un campo descriptivo mas amplio y además, a nivel espacial contienen patrones visuales que pueden guiar de una mejor manera la discriminación de lesiones csPCa. Sin embargo, estas características suelen ser numerosas y altamente variables. Para capturar patrones relevantes entre ellas suelen usarse descriptores compactos como lo son las matrices SPD. Sin embargo, para el procesamiento y clasificación de estos descriptores compactos SPD, se debe considerar la geometría de la variedad Riemanniana donde ellos pertenecen, por ende, las estrategias convencionales de aprendizaje profundo no resultan apropiadas. En este orden de ideas, en la literatura se han propuesto redes especializadas en el procesamiento de matrices SPD las cuales han demostrado su alto rendimiento en tareas de aprendizaje profundo y escenarios con datos limitados ⁴⁶⁵⁰⁴⁹.

El conjunto de embebidos SPD compactos fueron calculados a partir del extractor de características $Enc(\cdot)$ pre-entrenado mediante la estrategia contrastiva. Estas características representan activaciones profundas de las ultimas capas convolucionales por modalidad $Enc^{T2}(\mathbf{X}_{T2})$, $Enc^{ADC}(\mathbf{X}_{ADC})$, $Enc^{HBV}(\mathbf{X}_{HBV})$. Retornado así un banco de k características profundas de dimensión $\tilde{W} \times \tilde{H} \times \tilde{D}$ para cada parámetro de la bp-MRI \mathbf{F}_{T2} , \mathbf{F}_{ADC} , \mathbf{F}_{HBV} . Posteriormente, se concatenaron para formar un único banco de características profundas general de los tres parámetros de la bp-MRI, denotado por \mathcal{F} (ver figura 5.c). Este banco de características es re-organizado en una matriz rectangular \mathbf{C} , aplanando cada característica profunda de \mathcal{F} , consiguiendo así que cada fila de la matriz de \mathbf{C} sea una característica profunda aplanada, obteniendo $\mathcal{F} \in \mathbb{R}^{3k \times \tilde{W} \times \tilde{H} \times \tilde{D}}$ y de esta forma $\mathbf{C} \in \mathbb{R}^{3k \times \tilde{W}\tilde{H}\tilde{D}}$. Para el calculo del descriptor SPD, se incorpora una capa que denominamos *SPDpooling*, la cual codifica una matriz SPD de la siguiente manera:

$$\mathbf{M} = SPD_{pool}(\mathbf{C}) = \frac{1}{\tilde{W} \times \tilde{H} \times \tilde{D}} \mathbf{C} \mathbf{C}^\top. \quad (6)$$

De esta manera, el banco de características profundas es compactado en una matriz $\mathbf{M} \in Sym_{++}^{3k}$ (de dimensión $3k \times 3k$), donde cada entrada $\mathbf{M}(i, j)$ guarda relaciones entre pares de característica profundas intra- e inter- secuencia MRI mediante el computo del producto interno entre pares de características aplanadas. Al ser SPD, y dada su geometría, denominamos a estos descriptores, como embebidos geométricos. En base a esto, para la clasificación de estos embebidos en este trabajo se propusieron dos estrategias que serán descritas a continuación.

4.4.1. Clasificación de embebidos geométricos SPD proyectados en el espacio tangente ($Log(SPD)$). Se decidió como primera alternativa para la clasificación de estos embebidos geométricos SPD proyectarlos a un espacio Euclidiano previo a realizar su clasificación final. Para esto, se realizó la proyección de los embebidos al plano tangente en la identidad mediante la función logaritmo:

$$f_{LogEig}(\mathbf{M}) = Log(\mathbf{M}) = U Log(\Sigma) U^\top, \quad (7)$$

donde, $\mathbf{M} = U \Sigma U^\top$ es la descomposición espectral de \mathbf{M} , esto es, U es la matriz de autovectores y Σ una matriz diagonal con los autovalores $\lambda_1, \dots, \lambda_n$. Para el calculo de la función $Log(\mathbf{M})$, es necesario calcular el logaritmo de cada uno de los autovalores de \mathbf{M} .

$$\log(\Sigma) = \begin{bmatrix} \log(\lambda_1) & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \log(\lambda_n) \end{bmatrix}. \quad (8)$$

De esta manera, dada una matriz SPD \mathbf{M} , se obtiene una matriz $f_{LogEig}(\mathbf{M})$ que resulta proyectada en el plano tangente a la variedad de las matrices SPD, ver Figura 7. Una vez proyectada esta matriz SPD, dado que resulta simétrica, se tomó la parte triangular superior para prevenir repetición de características. Estos vectores embebidos, ahora euclidianos, junto

a sus respectivas etiquetas, serán clasificados para evaluar la capacidad de las representaciones SPD de resumir patrones discriminativos de las lesiones csPCa.

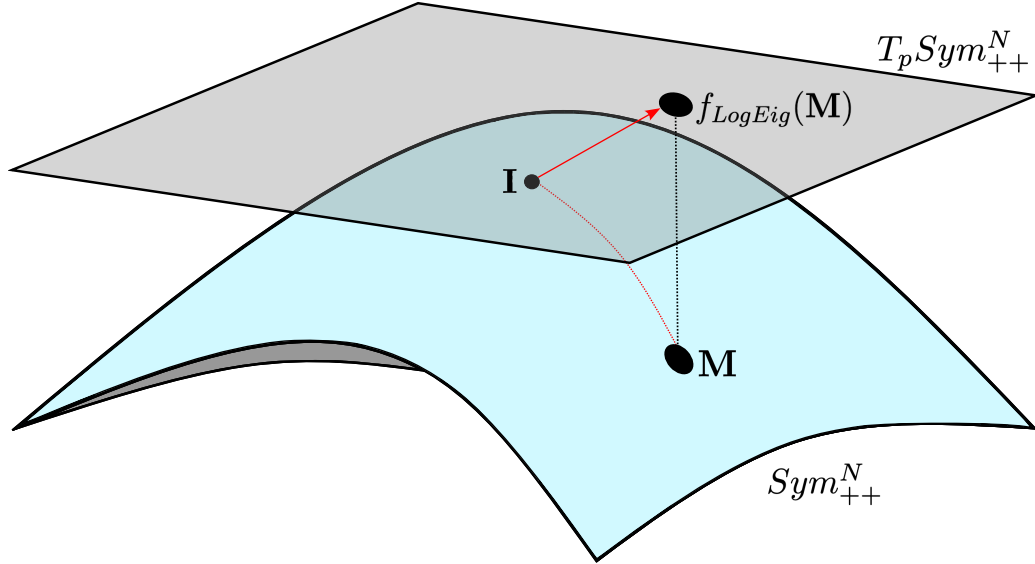


Figura 7

Mapeo Riemanniano de la función logaritmo. Considerando que $\mathbf{M} \in \text{Sym}_{++}^N$, se presenta el mapeo desde el espacio Riemanniano Sym_{++}^N al plano tangente de la matriz identidad \mathbf{I} , $T_p \text{Sym}_{++}^N$.

4.4.2. Clasificación mediante redes profundas SPD (SPDnet). En la literatura existen redes de aprendizaje profundo sobre matrices SPD, que permiten aprovechar de una mejor manera las propiedades geométricas de estas matrices, alcanzando representaciones más discriminativas. En esta parte del trabajo, se implementó la red de aprendizaje profundo sobre matrices SPD (SPDnet) planteada por *Huang et.al*, la cual propone dos capas *BiMap*, *ReEig*⁵⁰. A continuación se profundiza más en estas capas de aprendizaje.

- **BiMap:** En esta capa, a partir de una matriz SPD, se genera una nueva matriz de menor dimensión y mediante un mapeo bilineal $f_{BiMap}^{(\ell)}$:

$$\mathbf{M}_\ell = f_{BiMap}^{(\ell)}(\mathbf{M}_{\ell-1}, W_\ell) = W_\ell \mathbf{M}_{\ell-1} W_\ell^\top \quad (9)$$

Esta capa procesa como entrada una matriz SPD $\mathbf{M}_{\ell-1} \in \text{Sym}_{++}^{\ell-1}$, en función de una

matriz de pesos o de transformación entrenable $W_\ell \in \mathbb{R}_*^{d_\ell \times d_{\ell-1}}$ ($d_{\ell-1} > d_\ell$), donde $d_{\ell-1}$ es la dimensión de la matriz de entrada y d_ℓ es la dimensión de la matriz de salida. En esta matriz de pesos W_ℓ sus filas deben ser linealmente independientes para que la matriz resultante siga siendo SPD. Esta caracterización de los pesos hace necesario durante la optimización de estas matrices, considerar la variedad compacta de Stiefel $St(d_\ell, d_\ell - 1)$ donde pertenecen, aportando así mejor estabilidad numérica ⁵⁰⁴⁹.

- **ReEig:** Esta capa f_{ReEig}^ℓ de regularización no-lineal está inspirada en la función de activación *ReLU* de las redes neuronales profundas y se define de la siguiente manera:

$$\mathbf{M}_\ell = f_{ReEig}^{(\ell)}(\mathbf{M}_{\ell-1}) = U_{\ell-1} \max(\epsilon I, \Sigma_{\ell-1}) U_{\ell-1}^\top \quad (10)$$

donde:

$$\Sigma_{\ell-1} = \begin{bmatrix} \lambda_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \lambda_n \end{bmatrix} \quad (11)$$

Acá, se regulariza una matriz SPD de entrada utilizando su descomposición espectral $\mathbf{M}_{\ell-1} = U_{\ell-1} \Sigma_{\ell-1} U_{\ell-1}^\top$. Además, ϵ es el hiperparámetro de regularización, I es una matriz de identidad de $d_{\ell-1} \times d_{\ell-1}$, $U_{\ell-1}$ corresponde a una matriz cuyos elementos son los vectores-propios, $\Sigma_{\ell-1}$ es una matriz diagonal cuyos elementos son los valores-propios de $\mathbf{M}_{\ell-1}$ así como se ilustra en (11) y $\max(\epsilon I, \Sigma_{\ell-1})$ es la operación de rectificación de valores-propios de $\mathbf{M}_{\ell-1}$ donde a partir del hiperparámetro ϵ se analiza cada valor-propio reemplazando cada uno de ellos siempre y cuando sea menor que el umbral $\Sigma_{\ell-1}(i, i) < \epsilon$ esta operación se define de la siguiente manera:

$$A(i, i) = \begin{cases} \Sigma_{\ell-1}(i, i), & \Sigma_{\ell-1}(i, i) > \epsilon \\ \epsilon, & \Sigma_{\ell-1}(i, i) \leq \epsilon \end{cases} \quad (12)$$

Esta capa además de agregar no linealidad al modelo, contribuye a que los valores-propios de cada matriz no se devalúen, manteniendo la positividad de ellas, además de conservar la geometría de los embebidos ⁵⁰⁷⁰.

El uso de las capas *BiMap* y *ReEig* suele hacerse de manera secuencial, utilizando siempre una capa *ReEig* luego de una capa *BiMap*, esto se define como modulo Riemanniano (*BiRe Block*) el cual se puede ser concatenado múltiples veces. Finalmente, para realizar la clasificación en un espacio euclidiano la SPDnet concluye con una proyección al espacio tangente usando la función $f_{LogEig}^{(\ell)}(\cdot)$ definida en (7). Al realizar esta proyección, se toma el vector de la parte triangular superior, para posteriormente ser proyectado mediante una capa densa y una capa de clasificación (ver figura 5.c).

⁷⁰ P-A ABSIL; Robert MAHONY, and Rodolphe SEPULCHRE. *Optimization algorithms on matrix manifolds*. Princeton University Press, 2008.

5. DISEÑO EXPERIMENTAL

5.1. CONJUNTO DE DATOS

En este trabajo, se utilizó el conjunto de datos público del desafío *PI-CAI (Prostate Imaging: Cancer AI)*⁷¹⁷². *PI-CAI* proporciona en su totalidad 1.500 estudios BP-MRI, de los cuales 1.075 son benignos, 425 tienen lesión csPCa, sin embargo solamente 220 tienen marcación asociada a la lesión. Estas marcaciones están dadas por una delineación realizada por un radiólogo experto y tiene asociado como etiqueta su correspondiente grado histopatológico (ISUP 0-5) que revela la significancia clínica de la lesión.

Particularmente, en este trabajo se utilizó en total un conjunto de 1.295 estudios, de los cuales 1.075 corresponden a casos benignos (ISUP-0 e ISUP-1) y 220 casos con lesión csPCa (ISUP ≥ 2). En la Tabla 1, se detalla la cantidad de lesiones para cada clase, así como su respectiva grado de correspondencia con la escala ISUP. Adicionalmente, este conjunto de datos cuenta con un conjunto de variables clínicas asociadas a cada uno de los casos, tanto benignos como malignos. Entre las variables se cuenta con información de la Edad, el nivel de antígeno prostático específico (*PSA*), ver la Tabla 2.

⁷¹ Anindo SAHA et al. “The PI-CAI challenge: public training and development dataset”. In: *Zenodo, Jun* (2022).

⁷² Anindo SAHA et al. “Artificial intelligence and radiologists at prostate cancer detection in mri—the pi-cai challenge”. In: *Medical Imaging with Deep Learning, short paper track*. 2023.

Tipo	Clase	Cantidad
Benigno	ISUP 0	847
	ISUP 1	228
Maligno	ISUP 2	131
	ISUP 3	51
	ISUP 4	20
	ISUP 5	18

Tabla 1

Cantidad de estudios. En este estudio se consideraron 1.075 casos benignos, correspondientes a lesiones grado ISUP-0 e ISUP-1. En total se contó con 220 casos malignos con delineación de grado ISUP asociada, estos son aquellos con grado ISUP ≥ 2 .

Categoría	Edad Mediana (IQR)	PSA Mediana (IQR)
Benigno (ISUP 0-1)	65 (60-69)	7.85 (5.4-11.8)
Maligno (ISUP 2-5)	68 (63-75)	10.7 (7.6-18)

Tabla 2

Variables clínicas del conjunto de datos. Valor de la mediana y el rango intercuartílico (IQR) para las variables clínicas: Edad y concentración de antígeno prostático específico (PSA, unidades ng/ml) en test de sangre realizado al paciente. Cada valor se encuentra conglomerado para el conjunto de casos benigno y malignos.

En cuanto al preprocesamiento, el desafío *PI-CAI* proporciona estudios bp-MRI constituidos por las secuencias **T2W**, **ADC** y **DWI**, las cuales inicialmente cada una de ellas fue sometida a un remuestreo llevando todas las secuencias a un mismo espaciado de voxels de $0.5mm \times 0.5mm \times 3.0mm$, posteriormente se hizo un recorte al tamaño 384×384 con 24 cortes (*slices*)⁷². En este proceso, se aseguró que cada estudio **ADC** y **DWI** estuviera alineado con su correspondiente secuencia **T2W**. Por último cada secuencia por separado fue normalizada en el intervalo de $[0, 1]$. En el presente trabajo, no se usó como entrada la MRI en su totalidad, por el contrario, se definió una estrategia para extraer parches volumétricos para cada una de las secuencias (**T2W**, **ADC** y **DWI**), obteniendo una región de interés (ROI) tanto para casos con lesión y como para casos sin lesión. En los casos donde hay lesión csPCa, se extrajo una ROI volumétrica centrada en el centro de masa de la delineación dada del experto radiólogo.

Para los casos en los no existe lesión csPCa, se seleccionó un punto aleatoriamente dentro de la glándula prostática (p_x, p_y) dentro de un corte aleatorio p_z acotado en los slices [9, 15], para obtener (p_x, p_y, p_z) , ver figura 8. Bajo esta extracción de parches volumétricos, finalmente en este trabajo se usaron 1.075 ROIs benignas y 220 con lesión csPCa.

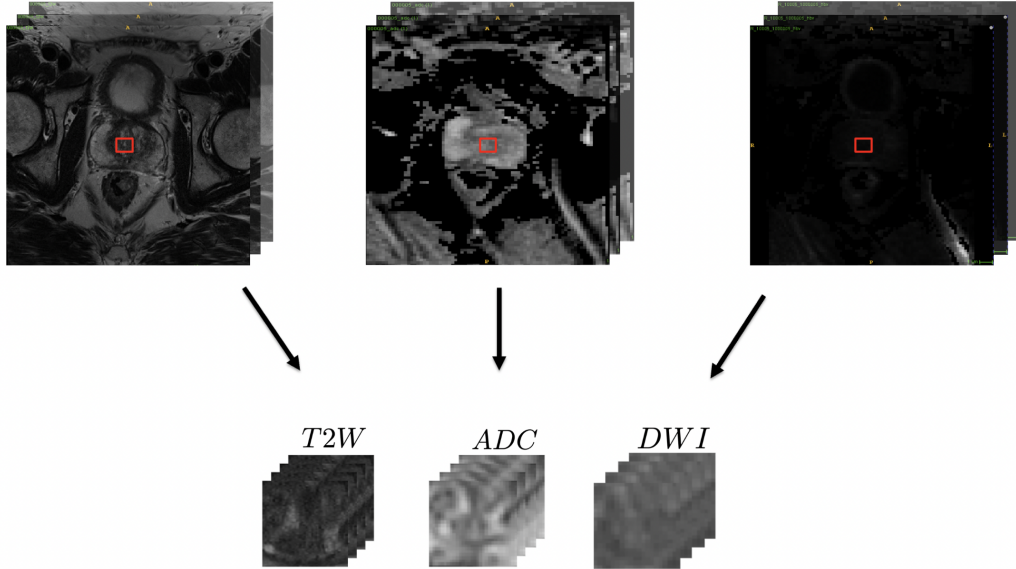


Figura 8
Parches volumétricos bp-MRI. Se presenta un ejemplo de una región de interés (ROI) volumétrica calculada en la misma cordenada para las diferentes secuencias bp-MRI

5.2. CONFIGURACIÓN EXPERIMENTAL

5.2.1. Arquitectura red encoder y proyección. El encoder $Enc(\cdot)$ implementado en este trabajo, está compuesto por una secuencia de bloques convolucionales tridimensionales diseñados para procesar parches volumétricos que representan estudios bp-MRI. El modelo comienza con una capa convolucional 3D con un kernel de tamaño $(1, 3, 3)$ y 8 filtros de salida, seguida de una capa de normalización por lotes y una función de activación ReLU. A continuación, se aplica otra convolución 3D con un kernel de tamaño $(3, 3, 3)$ y 8 filtros de salida, nuevamente seguida por normalización por lotes y activación ReLU. Después de este primer bloque convolucional, se emplea una capa de MaxPooling 3D con un kernel de tamaño $(1, 2, 2)$ para

reducir la resolución espacial de las características, manteniendo la profundidad. Este patrón se repite en bloques sucesivos, incrementando el número de filtros a 16, 32 y 64 en cada bloque subsiguiente, respectivamente, mientras que la dimensión espacial se reduce progresivamente a través de operaciones de MaxPooling. El encoder finaliza con una capa convolucional con 64 filtros de salida y un kernel de tamaño (1, 3, 3).

Seguido de $Enc(\cdot)$, la capa de proyección $Proy(\cdot)$ inicia con una capa lineal que recibe el vector resultado del encoder y lo proyecta en un vector embebido de dimensión 1024. Esta capa es seguida por una activación no lineal $LeakyReLU$. Posteriormente, otra capa lineal reduce las dimensiones a 256, seguida nuevamente por la activación $LeakyReLU$ con el mismo parámetro. Finalmente, una última capa lineal proyecta las características a un vector embebido de dimensión 128.

5.2.2. Esquema de aprendizaje contrastivo supervisado (SCL). Este enfoque representa nuestro modelo base (*Contrastivo*) que será evaluado y comparado posteriormente con los embebidos geométricos propuestos. Para el entrenamiento de esta arquitectura contrastiva se usaron 1500 épocas con tamaño de batch de 32. Para el ajuste se consideraron dos configuraciones: utilizando la pérdida $\mathcal{L}_{NT-Xent}$ y $\mathcal{L}_{Triplet}$ definidas en las Ecuaciones (4) y (5) respectivamente. Para la configuración de la pérdida $\mathcal{L}_{NT-Xent}$ se utilizó una temperatura $\tau = 0.07$ y como medida de similaridad se utilizó la similaridad del coseno. Para la configuración con la pérdida $\mathcal{L}_{Triplet}$, se usó un margen de $\alpha = 0.05$. En ambas configuraciones se utilizó la estrategia de *Early Stopping* con una paciencia de 75 épocas para evitar problemas de sobreaprendizaje. Como optimizador se eligió el RMSprop con una tasa de aprendizaje de 1×10^{-6} y momentum de 0.6. Adicionalmente, se utilizó la técnica de aumento de datos en línea durante el entrenamiento de la red mediante el uso de reflexiones horizontales, reflexiones verticales y traslaciones, implementadas de manera aleatoria. El enfoque propuesto fue entrenado y validado mediante una técnica de validación cruzada estratificada con 5 particiones (validación en *5k-fold*) proporcionada por el reto PI-CAI. En cada una de estas particiones se garantiza que el conjunto de entrenamiento y validación tengan una misma proporción entre casos csPCa

y benignos. Además se aseguró que las ROIs volumétricas extraídas de un mismo estudio MRI se mantuvieran en la misma partición, ya sea de entrenamiento o de validación.

5.2.3. Embebidos SPD. Los embebidos SPD fueron obtenidos a partir de las características profundas del encoder del modelo resultante del esquema contrastivo como se explicó en la Sección 4.4. Específicamente se tomaron 32 características de dimensión $6 \times 8 \times 8$ de la penúltima capa convolucional del encoder $Enc(\cdot)$ ajustado, para las secuencias T2, ADC y DWI. Con esto, se calculó así para cada ejemplo una respectiva matriz SPD de dimensión 96×96 . Estos descriptores SPD, fueron calculados con los datos de entrenamiento y validación con las respectivas particiones del *5k-fold* correspondiente.

5.2.4. Arquitectura SPDnet. La red de aprendizaje profundo geométrico SPDnet recibe como entrada inicial los embebidos SPD y está conformada por dos bloques *BiRe*, los cuales reducen la matriz \mathbf{M} de entrada de dimensión 96×96 a 48×48 y luego a 24×24 . El vector resultado de proyectar el resultado usando una capa *LogEig* y filtrar la parte triangular superior genera un vector de dimensión 300. Posteriormente, este vector es finalmente proyectado mediante una red densa a 64, 32 y 2 neuronas para su clasificación. Esta red SPDnet fue entrenada con la función de pérdida *CrossEntropyLoss* durante 80 épocas, con un tamaño de batch de 32, se usó el optimizador RMSprop con una tasa de aprendizaje de $\alpha = 0.0001$.

5.2.5. Clasificación de embebidos. A partir del esquema contrastivo, los siguientes métodos se utilizaron para finalmente clasificar representaciones aprendidas y extraídas de este modelo:

- **Método base (*Contrastivo*).** Los embebidos resultantes de la capa final de proyección del modelo contrastivo *Contrastivo*, fueron clasificados utilizando dos modelos de aprendizaje: Máquina de soporte vectorial (*Support Vector Machine SVM*) con un kernel RBF (*Radial basis function*) y una regresión logística (*Logistic Regression LR*) así como

se observa en la Figura 9.

- **Clasificación mediante proyección al plano tangente ($Log(SPD)$).** Se realizó la clasificación a partir del banco de matrices SPD de dimensión 96×96 . Para esto, las matrices fueron proyectadas al plano tangente mediante la función logaritmo, y vectorizadas, resultando en vectores de tamaño 4656. Para la clasificación se utilizaron los mismos modelos de aprendizaje del método *Contrastivo* (Ver Figura 9).
- **Clasificación utilizando la red profunda geométrica ($SPDnet$).** Utilizando los mismos descriptores compactos SPD del método $Log(SPD)$, se implementó como segundo método de clasificación geométrico una red SPDnet (Ver Figura 9).

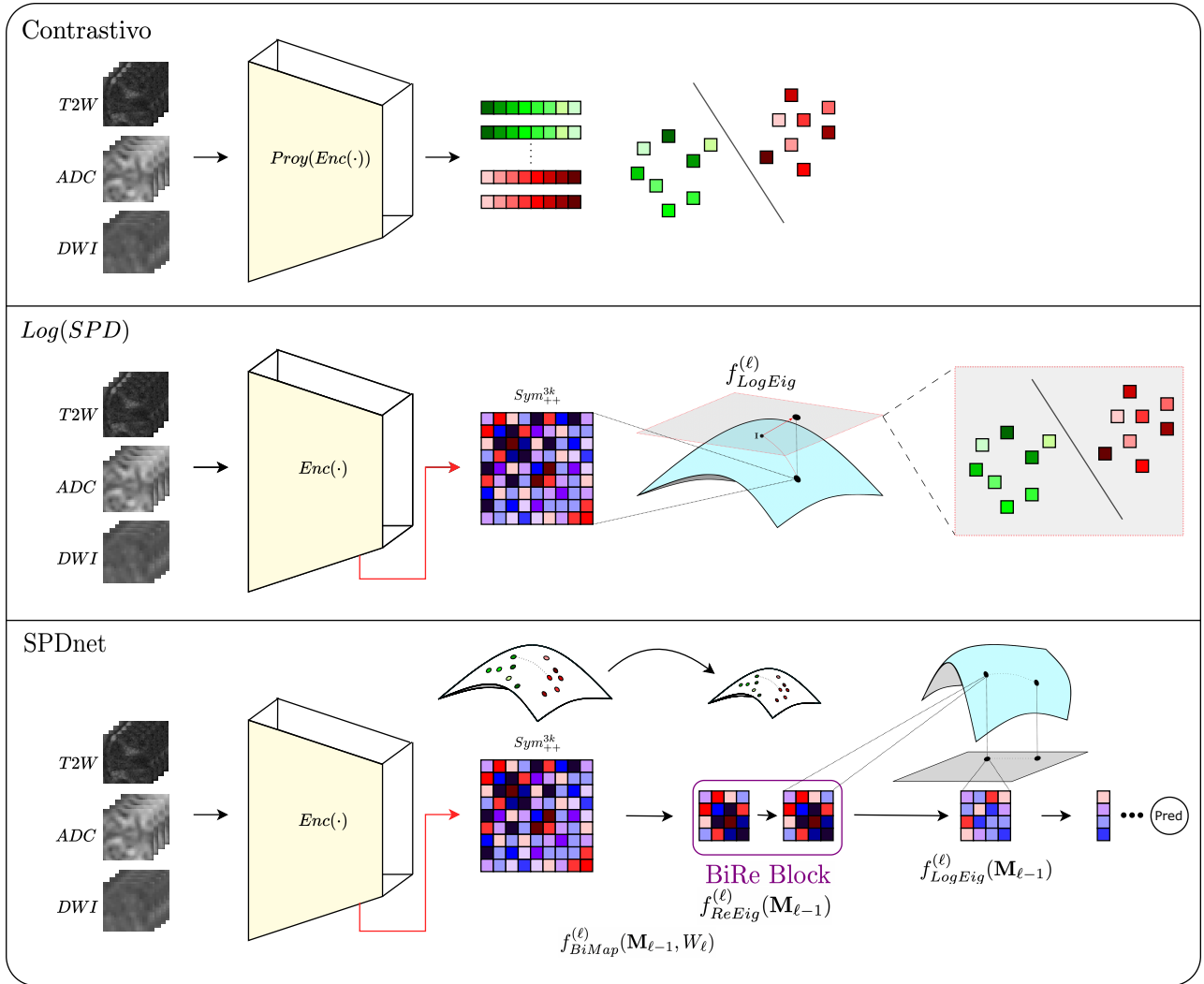


Figura 9

Métodos de clasificación de embebidos. La figura muestra los tres métodos de clasificación, donde *Contrastivo* es el método base que clasifica los embebidos obtenidos del modelo contrastivo, *Log(SPD)* es el método que para la clasificación usa los embebidos SPD proyectados al plano tangente y *SPDnet* es el método que a partir de los embebidos SPD, realiza aprendizaje Riemanniano para la clasificación.

Los embebidos usados en cada uno de estos métodos organizados con las 5 particiones proporcionadas, ajustando cada método con los correspondientes embebidos del subconjunto de entrenamiento y evaluados sobre los embebidos que corresponden a los datos de validación para cada partición.

5.3. MÉTRICAS

En este trabajo se utilizó como métricas de evaluación la exactitud (*Exactitud*), el área bajo la curva ROC (*AUC-ROC*), y el área bajo la curva de precisión-recall (*AUC-PR*). La exactitud calcula el porcentaje de las predicciones correctas respecto al total de predicciones, no obstante depende del umbral elegido para la clasificación, y además puede llegar a sesgar los resultados debido a la naturaleza del desbalance de clases del problema abordado. Por ello, adicionalmente se usó la métrica AUC-ROC, ya que que permite evaluar la capacidad del modelo para discriminar entre las clases positivas y negativas a través de diferentes umbrales, proporcionando una visión más completa para el análisis del rendimiento del modelo. En ultima instancia, al ser un problema donde cada falso positivo y falso negativo implica riesgos significativos para los pacientes, se usó el AUC-PR, el cual relaciona la precisión y la sensibilidad, permitiendo evaluar el rendimiento del modelo al clasificar las clases positivas.

En este trabajo estamos interesados en comparar espacios embebidos y medir su capacidad para discriminar entre lesiones csPCa benignas y malignas. Así, nosotros exploramos tres alternativas para generar los espacios embebidos: el espacio contrastivo, el espacio SPD tangente y el espacio SPD aprendido desde una representación geométrico. Para establecer la capacidad discriminatoria y el alcance de estos modelos, en este trabajo calculamos diferencias estadísticas significativas con respecto a las distribuciones de probabilidad de las predicciones de los modelos. Así pudimos establecer si los valores de separación de estos embebidos tenían significancia estadística. En este trabajo se implementó la prueba no paramétrica U de Mann-Whitney, que no tiene restricciones que los datos sigan una distribución normal, permitiendo evaluar el comportamiento de cada una de las representaciones en el clasificador.

6. EVALUACIÓN Y RESULTADOS

En este trabajo se exploraron representaciones embebidas en la tarea de discriminar lesiones de CsPa según su malignidad. Para ello se construyeron espacios geométricos de baja dimensionalidad, siguiendo ajustes contrastivos y codificaciones SPD. Para la validación de la metrología propuesta, inicialmente se realizó una validación sobre los tres métodos propuestos, analizando la capacidad para discriminar lesiones malignas, con respecto a las regiones control, según el estándar histopatológico. Todos los experimentos realizados se reportan bajo los conjuntos de validación del $5k$ -fold, reportando las métricas AUC-ROC, Exactitud y AUC-PR promedio en estas particiones. En este sentido, cada fold consideró 1.295 muestras (1.075 benignas y 220 malignas) para realizar el entrenamiento y evaluación.

Pérdida	Modelo	Clasificador	AUC-ROC \uparrow	Exactitud \uparrow	AUC-PR \uparrow
$\mathcal{L}_{NT-Xent}$	Contrastivo	LR	0.85 ± 0.04	0.88 ± 0.03	0.68 ± 0.10
		SVM	0.84 ± 0.05	0.61 ± 0.26	0.67 ± 0.10
	$Log(SPD)$	LR	0.92 ± 0.01	0.92 ± 0.01	0.82 ± 0.02
		SVM	0.93 ± 0.02	0.92 ± 0.01	0.83 ± 0.03
	SPDnet	-	0.92 ± 0.02	0.88 ± 0.04	0.79 ± 0.05
$\mathcal{L}_{Triplet}$	Contrastivo	LR	0.80 ± 0.04	0.84 ± 0.01	0.54 ± 0.13
		SVM	0.80 ± 0.04	0.77 ± 0.11	0.52 ± 0.13
	$Log(SPD)$	LR	0.92 ± 0.02	0.92 ± 0.01	0.81 ± 0.05
		SVM	0.93 ± 0.02	0.92 ± 0.01	0.83 ± 0.03
	SPDnet	-	0.91 ± 0.01	0.87 ± 0.06	0.75 ± 0.01

Tabla 3

Resultados de los métodos propuestos. Se presentan las métricas correspondientes a la clasificación de los métodos propuestos ajustados con las pérdidas $\mathcal{L}_{NT-Xent}$ y $\mathcal{L}_{Triplet}$. Para el método de clasificación de embebidos Contrastivo y el propuesto $Log(SPD)$ se consideraron dos clasificadores: SVM y LR. Además se evalúa el desempeño utilizando la red de aprendizaje profundo SPDnet. En **negrilla** se resaltan los mejores resultados para cada pérdida.

En el primer experimento se comparó el carácter discriminatorio de los tres conjuntos

de embebidos, generados desde el método estrictamente contrastivo (línea base del trabajo) y usando las configuraciones propuestas: $Log(SPD)$ y $SPDnet$. En la tabla 3 se ilustra los resultados obtenidos usando funciones de pérdida contrastivas: $\mathcal{L}_{NT-Xent}$ y $\mathcal{L}_{Triplet}$ y validando la clasificación con un regresor logístico (LR) y una máquina de soporte vectorial (SVM), usando un kernel RBF (*Radial basis function*). Como se puede observar, Las configuraciones geométricas alcanzaron un desempeño destacable, con respecto a la configuración contrastiva, alcanzando con una simple separación lineal una exactitud de 0.92 la composición geométrica con $log(SPD)$ y un AUC-ROC de 0.92, ajustando la representación con la pérdida $\mathcal{L}_{NT-Xent}$.

La estrategia de clasificación propuesta $Log(SPD)$, en conjunto con el clasificador SVM, fue la estrategia con mejor rendimiento, alcanzando un AUC-ROC de 0.93 y 0.92 de Exactitud, donde no hubo diferencia estadísticamente significativa al usar los embebidos calculados desde representaciones ajustadas con $\mathcal{L}_{NT-Xent}$ o $\mathcal{L}_{Triplet}$ ($p > 0.05$). Por otra parte, la segunda estrategia de clasificación de embebidos geométricos propuesta $SPDnet$, también obtuvo su mejor rendimiento con los embebidos ajustados desde la pérdida $\mathcal{L}_{NT-Xent}$, obteniendo un 0.92 de AUC-ROC y 0.88 de Exactitud, muy similar al rendimiento al usar la $\mathcal{L}_{Triplet}$, no encontrando diferencias significativas ($p > 0.05$). De esta forma, el mejor método fue $Log(SPD)$ para ambas pérdidas, siendo ligeramente superior que $SPDnet$, sin embargo, no se encontró diferencia estadística significativa en ninguno de los casos ($p > 0.05$).

A pesar de esto, observando la métrica AUC-PR, existe una amplia diferencia del método propuesto $Log(SPD)$ sobre la aproximación constrastiva (*Contrastivo*), con un AUC-PR de 0.83, 15% de ganancia en esta métrica, mientras que $SPDnet$ alcanzó un AUC-PR de 0.79, superando por un 11%. De esta manera se observó que los método propuestos para la clasificación desde embebidos geométricos SPD, superan ampliamente al modelo estrictamente contrastivo, resaltando la relevancia de considerar relaciones entre características profundas aprendidas en el enfoque contrastivo mediante el cálculo de los descriptores SPD.

Teniendo en cuenta la limitación en la disponibilidad de datos etiquetados en la rutina clínica, uno de nuestros objetivos es evaluar el potencial de las estrategias propuestas en estos

entornos con pocos datos. Espor ello, que en una segunda evaluación, se midió el rendimiento de los métodos propuestos utilizando el 20%, 40%, 60% y 80% de los datos de entrenamiento en cada una de las particiones del *5k-fold* proporcionado, manteniendo la misma cantidad de datos de validación, usando los mismos hiperparámetros y utilizando las mejores configuraciones de cada estrategia en la anterior evaluación, estas fueron: *Contrastivo* con *LR*, *Log(SPD)* con *SVM*, y *SPDnet*. Se consideraron tanto los casos ajustados con $\mathcal{L}_{NT-X_{ent}}$ como los casos ajustados con $\mathcal{L}_{Triplet}$. Los resultados se presentan en la Figura 10.

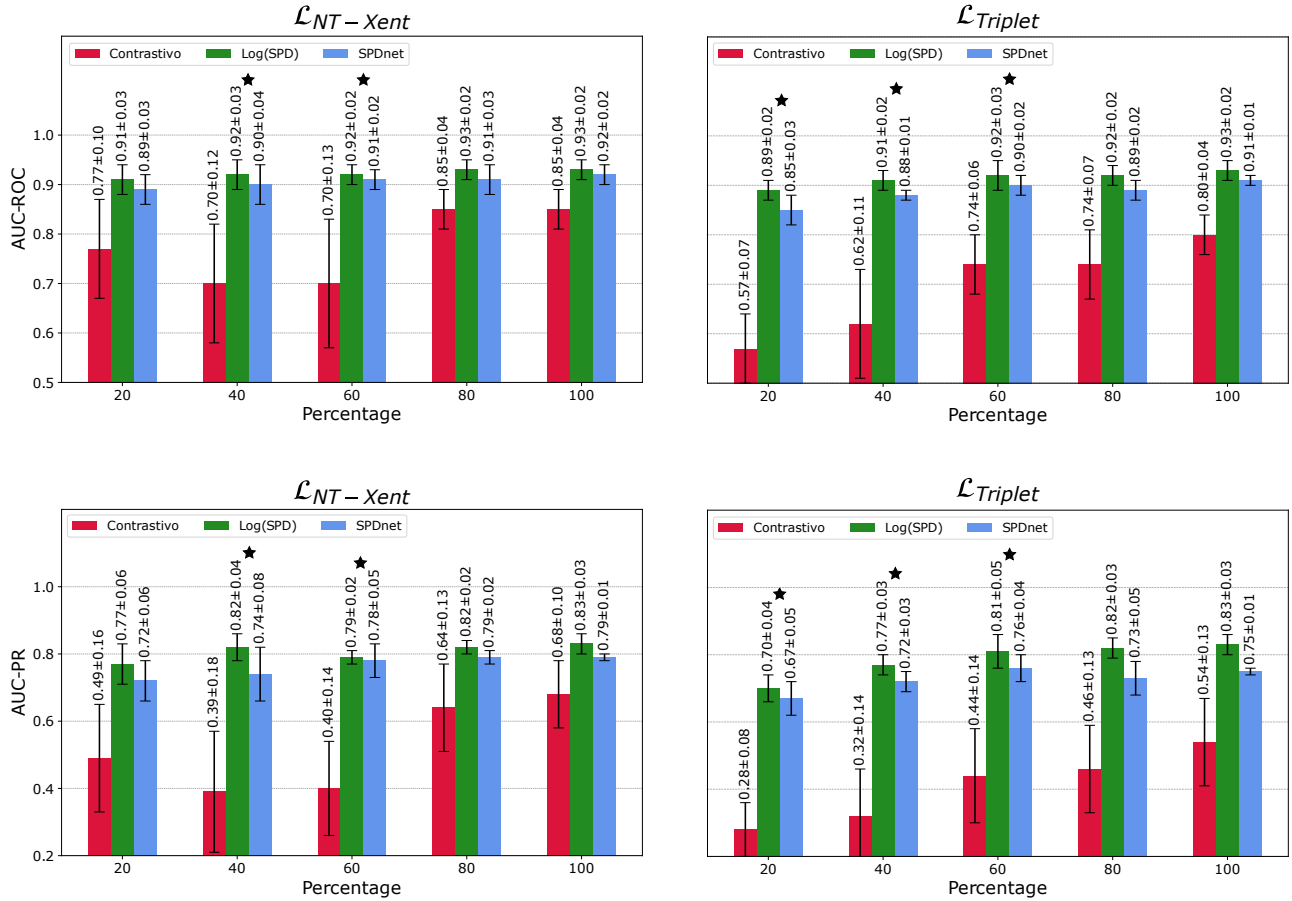


Figura 10

Rendimiento de los modelos en escenarios con pocos datos. Se presenta el rendimiento de los tres métodos propuestos para las métricas AUC-ROC y AUC-PR, utilizando distintos porcentajes del conjunto de datos de entrenamiento con las dos funciones de pérdida. Además se ilustra con una estrella (\star) en los casos que si hay diferencia estadísticamente significativa ($p < 0.001$) entre el método Log(SPD) y SPDnet

En la Figura 10 se observa que al usar los embebidos codificados sobre $\mathcal{L}_{NT-Xent}$, el *Contrastivo* fue el método menos robusto, obteniendo con todos los datos de entrenamiento para su ajuste un 0.85 de AUC-ROC y 0.68 de AUC-PR, mientras que al reducir los datos al 20% el rendimiento disminuyó hasta 0.77 y 0.49, presentando una disminución del 9.41% y 27.94% en AUC-ROC y AUC-PR respectivamente. Por otro lado, el método *SPDnet* mostró solamente una disminución del 3.26% y 8.86% en AUC-ROC y AUC-PR respectivamente al

pasar de usar el 100% de los datos de entrenamiento a solo el 20%. Por ultimo, el mejor método fue $Log(SPD)$ el cual reportó un 0.93 y un 0.91 de AUC-ROC usando el 100% y 20% de los datos de entrenamiento respectivamente, representando solo una disminución del 2.15% del rendimiento en esta métrica, con respecto al AUC-PR, el rendimiento decayó solamente un 8.86%. Esto demuestra que al incorporar los descriptores SPD como alternativa de codificación de representaciones embebidas se alcanza una mayor robustez en escenarios con pocos datos para el caso de las representaciones ajustadas desde $\mathcal{L}_{NT-Xent}$.

También se analizaron las representaciones aprendidas desde $\mathcal{L}_{Triplet}$ para evaluar la relevancia de la pérdida contrastiva seleccionada en escenarios con pocos datos. Aquí se observó que en general todos los modelos se desempeñaron peor al reducir los datos de entrenamiento, donde el método contrastivo (*Contrastivo*), $Log(SPD)$, y $SPDnet$ mostraron una pérdida en el desempeño de 28.75%, 4.30%, y 6.59% en AUC-ROC, y 48.14%, 15.66%, 10.66%, en AUC-PR al usar el 100% y 20% de los datos, respectivamente. De este análisis, se observa que la $\mathcal{L}_{Triplet}$ no es una pérdida muy robusta en comparación con la $\mathcal{L}_{NT-Xent}$ en la tarea de aprendizaje de representaciones en entornos con pocos datos. Particularmente, estos resultados resaltan la relevancia en la elección de la estrategia de minado de datos implementada, mostrando que aunque los descriptores geométricos brindan robustez en enfoques con pocos datos, también es importante la elección de la pérdida contrastiva seleccionada. La pérdida $\mathcal{L}_{NT-Xent}$ considera la métrica de similaridad del coseno para comparar representaciones de regiones de interés, lo cual se traduce en medir la similaridad entre estos vectores de información. Por el contrario, la pérdida $\mathcal{L}_{Triplet}$ usa la norma L2 para ajustar comparar representaciones. Esto sugiere que el mejor método alcanzado aprovecha de mejor manera el cálculo de similitud entre vectores, lo que en nuestro problema se traduce en efectivamente agrupar parches MRI que representen la misma clase y compartan patrones texturales, mientras aquellos de clase opuesta son disimilares, es decir, alejados.

Por otra parte, al reducir la cantidad de datos de entrenamiento, se observó que con la $\mathcal{L}_{NT-Xent}$ para el 40% y 60% las distribuciones de estos modelos son estadísticamente diferentes

($p < 0.001$). Así mostrando robustez de los métodos propuestos para tener un carácter discriminatorio, aprendiendo con un conjunto limitado de datos. También se evidenció lo mismo al usar $\mathcal{L}_{Triplet}$ en entornos con solo el 20% hasta el 60% de los datos ($p < 0.001$). Lo anterior indica que la codificación de representaciones embebidas como matrices SPD en el aprendizaje, proporcionan una mayor robustez en escenarios con limitación de datos etiquetados. Sin embargo, al contrario de lo que se esperaba, el método de aprendizaje profundo geométrico (*SPDnet*) fue ligeramente superado por el método que implementa una proyección al plano tangente para llevar a cabo la clasificación ($\text{Log}(SPD)$).

Con el fin de evaluar la potencial capacidad de los métodos propuestos para estratificar diversos niveles de malignidad de lesiones csPCa, en la Figura 11 se muestra la distribución de las predicciones de los tres mejores modelos a partir de la pérdida $\mathcal{L}_{NT-Xent}$. También, se analizó la distribución de las predicciones de los modelos sobre ejemplos de tejidos sanos. A pesar de que los modelos fueron entrenados con etiquetas binarias, este análisis permite explorar su potencial para graduar el nivel de malignidad de siguiendo como referencia el sistema de estratificación ISUP.

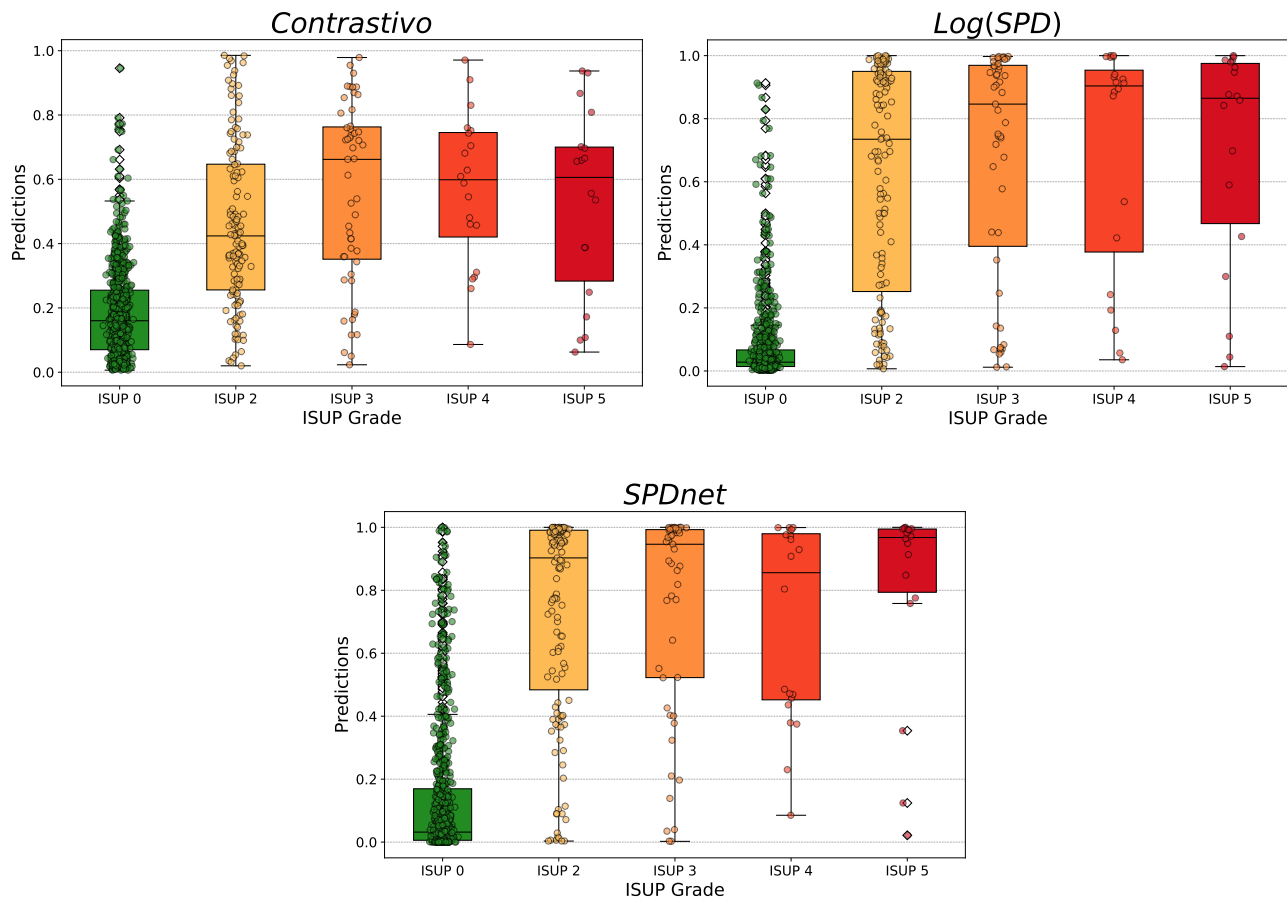


Figura 11

Distribución de predicciones en grados ISUP. Se ilustra la distribución de probabilidad de las predicciones de los métodos propuestos para casos sanos (ISUP 0) y distintos niveles de malignidad del cáncer de próstata (ISUP 2-5).

En la Figura 11 se observa que para el método contrastivo no hay una diferenciación clara entre los diferentes grados de malignidad, e incluso con respecto a las muestras sanas. Particularmente, a este método le cuesta discriminar las muestras ISUP 2. Contrario, para el método *Log(SPD)* se observó una mejor diferenciación de los casos csPCa con respecto a las muestras sanas, alcanzando unos valores de media de cada clase mas cercanos a 1, en comparación con el método *Contrastivo*. Sin embargo, persiste una diferenciación difícil entre las lesiones ISUP-2 y las muestras sanas. Por ultimo, para el caso del método *SPDnet*, a pesar de que

no fue el mejor respecto a las métricas reportadas, se observa que caracteriza considerablemente mejor los grados malignos, donde en general la media de las predicciones sobre cada grado ISUP, es mas alta que la del método $Log(SPD)$ correspondientemente, incluso para el grupo ISUP-2. Sin embargo, para muestras sanas, se observa una mayor dispersión y confusión hacia la clase positiva. Este análisis soporta el potencial del método propuesto para diferenciar lesiones malignas, especialmente de alta malignidad (alto grado ISUP). Por el contrario, se evidencia en diferencias las muestras correspondientes a ISUP-2.

7. CONCLUSIONES Y TRABAJO FUTURO

En este trabajo se desarrolló una estrategia de aprendizaje profundo para la clasificación de vectores embebidos entre lesiones csPCa y regiones benignas proyectadas desde una red multimodal bp-MRI. Para esto, inicialmente se adaptó un conjunto de datos bp-MRI con anotaciones asociadas a lesiones sanas y clínicamente significativas. A partir de esto, se implementó una estrategia de aprendizaje contrastivo supervisado para el ajuste de un espacio embebido de representaciones de regiones de interés MRI, buscando acercar embebidos de la misma clase y alejar ejemplos de clase opuesta. En este proceso, se involucró una red convolucional multimodal 3D sobre secuencias bp-MRI que permitió capturar patrones radiológicos de cada parámetro MRI de manera independiente, para luego fusionarlos y codificar los vectores embebidos. Adicionalmente, se utilizó una estrategia de minado de datos para la selección de los pares positivos y negativos entre embebidos, para un mejor ajuste del enfoque contrastivo.

Los embebidos alcanzados por la tarea de aprendizaje contrastivo, de manera tradicional son clasificados mediante algoritmos de clasificación de aprendizaje automático. Como alternativa, se propuso la codificación de representaciones embebidas geométricas a través de matrices simétricas definitas positivas (SPD), a partir de las activaciones convolucionales del modelo contrastivo pre-entrenado. Para la clasificación de los embebidos SPD, primero se realizó una estrategia de mapeo al plano tangente para ser clasificados mediante algoritmos de aprendizaje de máquina (método $\text{Log}(SPD)$). Por otra parte, se implementó una arquitectura de aprendizaje profundo geométrico que considera la variedad diferenciable donde están los embebidos SPD, aprendiendo a codificar embebidos geométricos más discriminativos, y finalmente llevando a cabo la clasificación de lesiones csPCa (método $SPDnet$). Se pudo observar que los enfoques geométricos propuestos $\text{Log}(SPD)$ y $SPDnet$ alcanzaron el mejor desempeño con AUC-ROC de 0.93 y 0.92 respectivamente, superando al método *Contrastivo* por 8% y 7%. Estos resultados demuestran la habilidad de estos descriptores SPD para codificar de manera apropiada variaciones intra- e inter- clase, comúnmente problemáticas en el contexto de imágenes MRI, para

capturar patrones relacionados con el Cáncer de Próstata. En un estudio con un número de datos limitado para el entrenamiento, se identificó que el método $Log(SPD)$ mostró un mejor desempeño con respecto a utilizar el *Contrastivo* o el método *SPDnet*, al usar el 60% o menos del conjunto de datos de entrenamiento. De esta manera, se identificó la ventaja de los embebidos SPD para caracterizar lesiones csPCa, considerando relaciones entre representaciones profundas radiológicas, incluso en entornos con pocos datos, desmejorando un 2% en comparación con el *Contrastivo* que desmejoró un 8% al pasar de usar el 100% al 20% de los datos de entrenamiento

En el estado del arte se han implementado diferentes estrategias para la clasificación de lesiones csPCa, utilizando el conjunto de datos público (*PROSTATEx*), el cual cuenta con 204 pacientes con 320 lesiones sospechosas y un conjunto de prueba de 140 pacientes con 200 lesiones. Por ejemplo, el trabajo de Mehrtash *et al.* propuso una arquitectura multimodal convolucional utilizando mp-MRI que incluía información zonal de la ubicación de la próstata. Este trabajo obtuvo un 0.80 de AUC-ROC ⁵². También Liu *et al.* propusieron una arquitectura convolucional 3D *end to end*, que utiliza MRI multiparamétrica, en este método se usa un alto aumento de datos, obteniendo en el mismo conjunto de datos prostatex, un puntaje de AUC-ROC 0.84 ⁵⁵. De igual manera, Chen *et al.* propusieron de manera alternativa utilizar una estrategia de aprendizaje por transferencia (*Transfer learning*) utilizando arquitecturas clásicas como *Inception V3* y *Vgg-16*, preentrenadas en ImageNet, bajo esta estrategia, reportaron un AUC-ROC de 0.83 en su mejor configuración ⁵³. Del mismo modo, Gutierrez *et al.* propusieron una arquitectura convolucional utilizando bp-MRI, donde la clasificación final está a cargo de una red SPDnet logrando un 0.85 de AUC-ROC, en el conjunto de datos prostatex ⁴⁶. De manera general, observamos que para estos métodos, el AUC-ROC varía desde 0.80 hasta 0.85 en el mejor de los casos. De manera sobresaliente, el método propuesto fue validado en PICA con (1075 muestras positivas y 220 muestras negativas), demostrando un 0.93 en AUC-ROC. A pesar de no ser implementados sobre exactamente el mismo conjuntos de datos, ni la misma partición de entrenamiento que el usado en este trabajo, podemos observar un comportamiento significativo, validándose en escenarios incluso con pocas muestras para el entrenamiento.

El trabajo propuesto presenta algunas limitaciones que podrían ser abordadas en investigaciones futuras. En primer lugar, la estrategia propuesta no es completamente entrenable "End to End", ya que la clasificación de los embebidos se realizó en una etapa posterior al ajuste contrastivo, lo que podría optimizarse y realizarse de manera conjunta en futuras aproximaciones. También, la selección de los parches benignos empleados en la tarea contrastiva podría mejorarse, por ejemplo, considerando la distribución natural de las lesiones malignas, lo que permitiría generar una tarea contrastiva más compleja al considerar pares negativos más difíciles visualmente de diferenciar. Esto a su vez mejoraría la diferenciación entre patrones benignos y malignos. Por otro lado, es conocido que los enfoques contrastivos se benefician aún en mayor medida de estrategias de aprendizaje no supervisadas, lo que sugiere que las representaciones geométricas propuestas podrían ser el punto de partida de nuevos enfoques no supervisados, aprovechando aún mejor el espacio de las representaciones ajustadas.

Al contrario de lo esperado inicialmente, el enfoque de aprendizaje profundo *SPDnet* no fue el de mejor desempeño, por lo que trabajos futuros pueden involucrar nuevas estrategias de aprendizaje geométrico que aprovechen de mejor manera los descriptores compactos SPD de entrada. Por otra parte, para extender la caracterización de lesiones propuesta a un enfoque con distintos grados de malignidad, se encontró que los enfoques propuestos deben ser replanteados, considerando las dificultades que tuvieron para diferenciar las muestras menos agresivas (ISUP-2 e ISUP-3) de las muestras benignas. Esto resulta de gran valor en la práctica clínica, donde es más relevante asignar un nivel de malignidad a cada lesión, pues esto permite evaluar la prognosis y tratamiento futuro del paciente. Los métodos geométricos propuestos demostraron ser potencialmente valiosos en la práctica clínica para identificar regiones malignas. Su potencial aumentaría si se integraran con herramientas de localización, permitiendo un análisis de estudios completos MRI. Dada la escasez de datos, esta herramienta sería especialmente útil en entornos médicos con recursos limitados.

BIBLIOGRAFÍA

ABSIL, P-A; MAHONY, Robert, and SEPULCHRE, Rodolphe. *Optimization algorithms on matrix manifolds*. Princeton University Press, 2008 (cit. on p. 45).

ALDOJ, Nader et al. “Semi-automatic classification of prostate cancer on multi-parametric MR imaging using a multi-channel 3D convolutional neural network”. In: *European radiology* 30.2 (2020), pp. 1243–1253 (cit. on p. 29).

ARSIGNY, Vincent et al. “Geometric means in a novel vector space structure on symmetric positive-definite matrices”. In: *SIAM journal on matrix analysis and applications* 29.1 (2007), pp. 328–347 (cit. on pp. 28, 41, 44).

BAMMER, Roland. “Basic principles of diffusion-weighted imaging”. In: *European journal of radiology* 45.3 (2003), pp. 169–184 (cit. on p. 22).

BERMAN, Rose M et al. “DCE MRI of prostate cancer”. In: *Abdominal radiology* 41 (2016), pp. 844–853 (cit. on p. 13).

BEYER, Thomas et al. “PI-RADS 2.1–Image Interpretation: The Most Important Updates and Their Clinical Implications”. In: *RöFo-Fortschritte auf dem Gebiet der Röntgenstrahlen und der bildgebenden Verfahren*. Vol. 193. 07. Georg Thieme Verlag KG. 2021, pp. 787–796 (cit. on p. 18).

BOOTHBY, William M. *An introduction to differentiable manifolds and Riemannian geometry*. Academic press, 1986 (cit. on p. 28).

BRAY, Freddie et al. “Global cancer statistics 2022: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries”. In: *CA: a cancer journal for clinicians* 74.3 (2024), pp. 229–263 (cit. on p. 12).

CAGLIC, Iztok and BARRETT, Tristan. “Diffusion-weighted imaging (DWI) in lymph node staging for prostate cancer”. In: *Translational Andrology and Urology* 7.5 (2018), p. 814 (cit. on p. 21).

CHANDARANA, Hersh and TAOULI, Bachir. “Diffusion and perfusion imaging of the liver”. In: *European journal of radiology* 76.3 (2010), pp. 348–358 (cit. on p. 22).

CHEN, Haochen et al. “A tutorial on network embeddings”. In: *arXiv preprint arXiv:1808.02590* (2018) (cit. on p. 24).

CHEN, Quan et al. “A transfer learning approach for classification of clinical significant prostate cancers from mpMRI scans”. In: *Medical imaging 2017: Computer-aided diagnosis*. Vol. 10134. SPIE. 2017, pp. 1154–1157 (cit. on pp. 29, 63).

CHEN, Ting et al. “A simple framework for contrastive learning of visual representations”. In: *International conference on machine learning*. PMLR. 2020, pp. 1597–1607 (cit. on pp. 26, 27, 36, 38).

CUZICK, Jack et al. “Prevention and early detection of prostate cancer”. In: *The lancet oncology* 15.11 (2014), e484–e492 (cit. on p. 12).

DE VENTE, Coen et al. “Deep learning regression for prostate cancer detection and grading in bi-parametric MRI”. In: *IEEE Transactions on Biomedical Engineering* 68.2 (2020), pp. 374–383 (cit. on p. 30).

DESVIAT, Pilar Vallejo et al. “Difusión en resonancia magnética en el estudio de lesiones hepáticas”. In: *Cirugía Española* 91.1 (2013), pp. 9–16 (cit. on pp. 21, 23).

DIETRICH, Olaf et al. “Technical aspects of MR diffusion imaging of the body”. In: *European journal of radiology* 76.3 (2010), pp. 314–322 (cit. on p. 20).

DU, Xiangtong et al. “NRD-Net: a noise-resistant distillation network for accurate diagnosis of prostate cancer with bi-parametric MRI images”. In: *Multimedia Tools and Applications* 83.11 (2024), pp. 33597–33614 (cit. on p. 30).

FALANA, William Olurotimi; SERENER, Ali, and SERTE, Sertan. “Clinically Significant Prostate Cancer Classification Using Anatomical T2W MRI Sequences and Deep Learning”. In: *2023 7th International Symposium on Multidisciplinary Studies and Innovative Technologies (ISMSIT)*. IEEE. 2023, pp. 1–6 (cit. on p. 29).

FERLAY, Jacques et al. “Estimating the global cancer incidence and mortality in 2018: GLOBOCAN sources and methods”. In: *International journal of cancer* 144.8 (2019), pp. 1941–1953 (cit. on pp. 12, 32).

FERRARO, Simona; BUSSETTI, Marco, and PANTEGHINI, Mauro. “Serum prostate-specific antigen testing for early detection of prostate cancer: managing the gap between clinical and laboratory practice”. In: *Clinical Chemistry* 67.4 (2021), pp. 602–609 (cit. on p. 16).

GE, Weifeng. “Deep metric learning with hierarchical triplet loss”. In: *Proceedings of the European conference on computer vision (ECCV)*. 2018, pp. 269–285 (cit. on p. 39).

GUTIÉRREZ, Yesid; AREVALO, John, and MARTÁNEZ, Fabio. “Multimodal Contrastive Supervised Learning to Classify Clinical Significance MRI Regions on Prostate Cancer”. In: *2022 44th Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC)*. IEEE. 2022, pp. 1682–1685 (cit. on pp. 25, 27, 31).

GUTIÉRREZ, Yesid; AREVALO, John, and MARTÍNEZ, Fabio. “A contrastive weakly supervised learning to characterize malignant prostate lesions in BP-MRI”. In: *Biomedical Signal Processing and Control* 96 (2024), p. 106584 (cit. on p. 25).

GUTIÉRREZ, Yesid et al. “A Multimodal Geometric Deep Representation to Support Bi-Parametric Prostate Cancer Lesion Classification”. In: *2023 IEEE 20th International Symposium on Biomedical Imaging (ISBI)*. IEEE. 2023, pp. 1–4 (cit. on pp. 27, 31, 41, 63).

HERMANS, Alexander; BEYER, Lucas, and LEIBE, Bastian. “In defense of the triplet loss for person re-identification”. In: *arXiv preprint arXiv:1703.07737* (2017) (cit. on p. 39).

HUANG, Zhiwu and VAN GOOL, Luc. “A riemannian network for spd matrix learning”. In: *Thirty-first AAAI conference on artificial intelligence*. 2017 (cit. on pp. 28, 41, 43–45).

JOHNSON, Linda M et al. “Multiparametric MRI in prostate cancer management”. In: *Nature reviews Clinical oncology* 11.6 (2014), pp. 346–353 (cit. on p. 18).

KEBAILI, Aghiles; LAPUYADE-LAHORGUE, Jérôme, and RUAN, Su. “Deep learning approaches for data augmentation in medical imaging: a review”. In: *Journal of Imaging* 9.4 (2023), p. 81 (cit. on p. 14).

KELLOFF, Gary J; CHOYKE, Peter, and COFFEY, Donald S. “Challenges in clinical prostate cancer: role of imaging”. In: *American journal of roentgenology* 192.6 (2009), pp. 1455–1470 (cit. on p. 19).

KHOSLA, Prannay et al. “Supervised contrastive learning”. In: *Advances in neural information processing systems* 33 (2020), pp. 18661–18673 (cit. on p. 26).

KLEINBAUM, David G et al. *Logistic regression*. Springer, 2002 (cit. on p. 40).

KOH, Dow-Mu and COLLINS, David J. “Diffusion-weighted MRI in the body: applications and challenges in oncology”. In: *American Journal of Roentgenology* 188.6 (2007), pp. 1622–1635 (cit. on p. 20).

LAHOTI, Amol Madanlal et al. “Role and application of diffusion-weighted imaging in evaluation of prostate cancer”. In: *Indian Journal of Medical and Paediatric Oncology* 39.03 (2018), pp. 349–354 (cit. on p. 23).

LAVALLEY, Michael P. “Logistic regression”. In: *Circulation* 117.18 (2008), pp. 2395–2399 (cit. on p. 40).

LAWRENTSCHUK, Nathan et al. “‘Prostatic evasive anterior tumours’: the role of magnetic resonance imaging”. In: *BJU international* 105.9 (2010), pp. 1231–1236 (cit. on p. 19).

LE-KHAC, Phuc H; HEALY, Graham, and SMEATON, Alan F. “Contrastive representation learning: A framework and review”. In: *Ieee Access* 8 (2020), pp. 193907–193934 (cit. on p. 24).

LIU, Qi; KUSNER, Matt J, and BLUNSOM, Phil. “A survey on contextual embeddings”. In: *arXiv preprint arXiv:2003.07278* (2020) (cit. on p. 24).

LIU, S et al. “Prostate Cancer Diagnosis using Deep Learning with 3D Multiparametric MRI. SPIE Med”. In: *Imaging* 10134 (2017), pp. 1–4 (cit. on pp. 29, 63).

MATA, Lidia Alcalá et al. “Artificial intelligence–assisted prostate cancer diagnosis: Radiologic–pathologic correlation”. In: *RadioGraphics* 41.6 (2021), pp. 1676–1697 (cit. on p. 14).

MEHRTASH, Alireza et al. “Classification of clinical significance of MRI prostate findings using 3D convolutional neural networks”. In: *Medical imaging 2017: computer-aided diagnosis*. Vol. 10134. SPIE. 2017, pp. 589–592 (cit. on pp. 29, 63).

MURPHY, Gillian et al. “The expanding role of MRI in prostate cancer”. In: *AJR Am J Roentgenol* 201.6 (2013), pp. 1229–1238 (cit. on pp. 12, 13, 18, 19).

NAJI, Leen et al. “Digital rectal examination for prostate cancer screening in primary care: a systematic review and meta-analysis”. In: *The Annals of Family Medicine* 16.2 (2018), pp. 149–154 (cit. on pp. 12, 13, 17, 32).

NAM, Robert K et al. “Increasing hospital admission rates for urological complications after transrectal ultrasound guided prostate biopsy”. In: *The Journal of urology* 183.3 (2010), pp. 963–969 (cit. on p. 17).

NOBLE, William S. “What is a support vector machine?” In: *Nature biotechnology* 24.12 (2006), pp. 1565–1567 (cit. on p. 40).

OCANTOS, Jorge A; PIETRANI, Marcelo Andrés, and PAGANINI, Lisandro. “Resonancia magnética de próstata: morfología y metabolismo”. In: *Revista Argentina de Urología* 72.4 (2007), p. 193 (cit. on p. 19).

OKOTIE, Onisuru T et al. “Characteristics of prostate cancer detected by digital rectal examination only”. In: *Urology* 70.6 (2007), pp. 1117–1120 (cit. on p. 16).

PATEL, Amit R and KLEIN, Eric A. “Risk factors for prostate cancer”. In: *Nature clinical practice Urology* 6.2 (2009), pp. 87–95 (cit. on p. 16).

RAJA, J et al. “Current status of transrectal ultrasound-guided prostate biopsy in the diagnosis of prostate cancer”. In: *Clinical radiology* 61.2 (2006), pp. 142–153 (cit. on p. 17).

REDA, Islam et al. “Deep learning role in early diagnosis of prostate cancer”. In: *Technology in cancer research & treatment* 17 (2018), p. 1533034618775530 (cit. on p. 28).

REMEIJER, Peter et al. “3-D portal image analysis in clinical practice: an evaluation of 2-D and 3-D analysis techniques as applied to 30 prostate cancer patients”. In: *International Journal of Radiation Oncology* Biology* Physics* 46.5 (2000), pp. 1281–1290 (cit. on p. 30).

RODRIGUEZ, Larissa V and TERRIS, Martha K. “Risks and complications of transrectal ultrasound guided prostate needle biopsy: a prospective study and review of the literature”. In: *The Journal of urology* 160.6 Part 1 (1998), pp. 2115–2120 (cit. on p. 17).

ROUVIÈRE, Olivier et al. “Artificial intelligence algorithms aimed at characterizing or detecting prostate cancer on MRI: How accurate are they when tested on independent cohorts?—a systematic review”. In: *Diagnostic and interventional imaging* 104.5 (2023), pp. 221–234 (cit. on p. 14).

SAHA, Anindo; HOSSEINZADEH, Matin, and HUISMAN, Henkjan. “End-to-end prostate cancer detection in bpMRI via 3D CNNs: effects of attention mechanisms, clinical priori and decoupled false positive reduction”. In: *Medical image analysis* 73 (2021), p. 102155 (cit. on p. 30).

SAHA, Anindo et al. “Artificial intelligence and radiologists at prostate cancer detection in mri—the pi-cai challenge”. In: *Medical Imaging with Deep Learning, short paper track. 2023* (cit. on pp. 46, 47).

SAHA, Anindo et al. “The PI-CAI challenge: public training and development dataset”. In: *Zenodo, Jun* (2022) (cit. on p. 46).

SALKA, Bassel R et al. “Effect of prostate MRI interpretation experience on PPV using PI-RADS version 2: a 6-year assessment among eight fellowship-trained radiologists”. In: *AJR. American journal of roentgenology* 219.3 (2022), p. 453 (cit. on p. 14).

SÁNCHEZ-ORO, R et al. “Resonancia magnética de próstata: guía práctica de interpretación e informe según PI-RADS versión 2.1”. In: *Radiología* 62.6 (2020), pp. 437–451 (cit. on p. 19).

SCIALPI, Michele et al. “Biparametric MRI of the prostate”. In: *Turkish journal of urology* 43.4 (2017), p. 401 (cit. on p. 18).

SHERRER, Rachael L et al. “Comparison of biparametric MRI to full multiparametric MRI for detection of clinically significant prostate cancer”. In: *Prostate cancer and prostatic diseases* 22.2 (2019), pp. 331–336 (cit. on p. 13).

SHOAG, Jonathan and BARBIERI, Christopher E. “Clinical variability and molecular heterogeneity in prostate cancer”. In: *Asian journal of andrology* 18.4 (2016), p. 543 (cit. on pp. 14, 32).

SHUKLA-DAVE, Amita and HRICAK, Hedvig. “Role of MRI in prostate cancer detection”. In: *NMR in Biomedicine* 27.1 (2014), pp. 16–24 (cit. on p. 17).

SUNG, Hyuna et al. “Global cancer statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries”. In: *CA: a cancer journal for clinicians* 71.3 (2021), pp. 209–249 (cit. on p. 32).

TAN, Cher Heng et al. “Dynamic contrast-enhanced MRI for the detection of prostate cancer: meta-analysis”. In: *American Journal of Roentgenology* 204.4 (2015), W439–W448 (cit. on p. 13).

THESTRUP, Karen Cecilie Duus et al. “Biparametric versus multiparametric MRI in the diagnosis of prostate cancer”. In: *Acta radiologica open* 5.8 (2016), p. 2058460116663046 (cit. on p. 18).

TUZEL, Oncel; PORIKLI, Fatih, and MEER, Peter. “Pedestrian detection via classification on riemannian manifolds”. In: *IEEE transactions on pattern analysis and machine intelligence* 30.10 (2008), pp. 1713–1727 (cit. on p. 28).

WANG, Peng et al. “Contrastive learning based hybrid networks for long-tailed image classification”. In: *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*. 2021, pp. 943–952 (cit. on p. 27).

WEISSER, Cedric et al. “Weakly supervised MRI slice-level deep learning classification of prostate cancer approximates full voxel-and slice-level annotation: Effect of increasing training set size”. In: *Journal of Magnetic Resonance Imaging* 59.4 (2024), pp. 1409–1422 (cit. on p. 29).

WU, Yawen et al. “Distributed contrastive learning for medical image segmentation”. In: *Medical Image Analysis* 81 (2022), p. 102564 (cit. on p. 27).

XINYANG, Song et al. “A machine learning radiomics model based on bpMRI to predict bone metastasis in newly diagnosed prostate cancer patients.” In: *Magnetic Resonance Imaging* 107 (2024), pp. 15–23 (cit. on p. 30).

XUAN, Hong; STYLIANOU, Abby, and PLESS, Robert. “Improved embeddings with easy positive triplet mining”. In: *Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision*. 2020, pp. 2474–2482 (cit. on pp. 37, 38).

ZHANG, Yuhao et al. “Contrastive learning of medical visual representations from paired images and text”. In: *Machine Learning for Healthcare Conference*. PMLR. 2022, pp. 2–25 (cit. on p. 27).