

SEGMENTACIÓN DE LESIONES DE CÁNCER DE PRÓSTATA EN SECUENCIAS DE
RESONANCIA MAGNÉTICA BIPARAMÉTRICA UTILIZANDO MECANISMOS DE
ATENCIÓN DE SEGUNDO ORDEN

MATEO GERÓNIMO ORTIZ CRUZATE

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

2025

SEGMENTACIÓN DE LESIONES DE CÁNCER DE PRÓSTATA EN SECUENCIAS DE
RESONANCIA MAGNÉTICA BIPARAMÉTRICA UTILIZANDO MECANISMOS DE
ATENCIÓN DE SEGUNDO ORDEN

MATEO GERÓNIMO ORTIZ CRUZATE

Trabajo de grado para optar por el título de:
Ingeniera de Sistemas

Director:

Fabio Martínez Carrillo

Doctor en Ingeniería de Sistemas y Computación

Codirector:

Juan Andrés Olmos Rojas

Magíster en Matemática Aplicada

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

2025

AGRADECIMIENTOS

En primer lugar, quiero expresar mi más profundo agradecimiento a mis padres, quienes con incontables noches de esfuerzo y desvelo me brindaron la oportunidad de formarme como profesional, un camino que ellos no tuvieron la fortuna de recorrer, pero que hicieron posible para mí con su entrega, sacrificio y dedicación. Este logro no es solamente mío, sino también de ustedes.

A mis hermanos, quienes me acompañaron por gran parte de este proceso, complementando los valores que me fueron inculcados, y que con su ejemplo de vida me enseñaron a actuar con integridad y responsabilidad dentro de la sociedad.

A mi director de tesis, Fabio Martínez, quien con su rigurosidad, orientación y paciencia me guió a lo largo de este fructífero proyecto. A todos los compañeros del grupo de investigación BIVL²ab, quienes con sus valiosos comentarios, críticas constructivas y aportes intelectuales contribuyeron a la consolidación de este trabajo.

A mi codirector de tesis, Juan Olmos, quien me acogió con paciencia y generosidad aun cuando no siempre fui el mejor estudiante. Su confianza en mis capacidades me motivó a seguir adelante. Más que un guía académico, fue un ejemplo de dedicación y compromiso, así como una persona valiosa cuya influencia trascendió lo estrictamente académico.

A mis amigos, quienes se convirtieron en una familia durante este camino. Con ellos compartí momentos de alegría y compañerismo que hicieron más llevadera esta etapa, y con su apoyo, orientación y compañía constante aportaron de manera significativa a este proceso.

CONTENIDO

	pág.
INTRODUCCIÓN	10
1. FUNDAMENTOS Y TRABAJOS PREVIOS	13
1.1. CÁNCER DE PRÓSTATA Y SECUENCIAS MRI	13
1.1.1 Secuencia T2 ponderada	16
1.1.2 Zonas prostáticas	18
1.2. ARQUITECTURAS PARA LA SEGMENTACIÓN DE IMÁGENES MÉDICAS . .	19
1.2.1 Arquitectura U-Net	19
1.2.2 Método de segmentación nnU-Net	21
1.3. ESTRATEGIAS COMPUTACIONALES PARA LA SEGMENTACIÓN DE LESIONES CSPCA	22
2. PROBLEMA DE INVESTIGACIÓN	26
3. OBJETIVOS	27
3.1. OBJETIVO GENERAL	27
3.2. OBJETIVOS ESPECÍFICOS	27
4. MÉTODO PROPUESTO	28
4.1. REPRESENTACIÓN BASADA EN U-NET	29
4.2. MECANISMOS DE ATENCIÓN Y APRENDIZAJE DE DESCRIPTORES DE SEGUNDO ORDEN	30
4.3. MÓDULO DE ATENCIÓN GEOMÉTRICA DE SEGUNDO ORDEN (SOGA) . .	33
4.4. DISEÑO EXPERIMENTAL	34
4.4.1 Datos.	34

4.4.2 Configuración de las arquitecturas	37
5. EVALUACIÓN Y RESULTADOS	41
5.1. DETECCIÓN EN LA COHORTE PÚBLICA DE PI-CAI	41
5.2. EVALUACIÓN DE GENERALIZACIÓN	43
6. CONCLUSIONES Y TRABAJO FUTURO	49
BIBLIOGRAFÍA	51
ANEXOS	59

LISTA DE FIGURAS

	pág.
Figura 1. Ejemplos de hallazgos en estudios bp-MRI	20
Figura 2. Arquitectura U-Net para la segmentación de lesiones csPCa en bp-MRI .	25
Figura 3. Figura del método propuesto SOGA	28
Figura 4. Diagrama de flujo de los conjuntos de datos	35
Figura 5. Predicciones en Prostate158	47

LISTA DE TABLAS

	pág.
Tabla 1. Desempeño en la cohorte pública PI-CAI con U-Net	41
Tabla 2. Desempeño en la cohorte pública PI-CAI con nnU-Net	42
Tabla 3. Desempeño en la cohorte PI-CAI Hidden Tuning	43
Tabla 4. Desempeño en el conjunto de datos Prostate158	44
Tabla 5. Desempeño según diferentes tamaños de lesión en Prostate158	46
Tabla 6. Resultados en zona central (CZ) y zona periférica (PZ)	48

RESUMEN

TÍTULO: Segmentación de lesiones de cáncer de próstata en secuencias de resonancia magnética bi-paramétrica utilizando mecanismos de atención de segundo orden. *

AUTOR: Mateo Gerónimo Ortiz Cruzate **

PALABRAS CLAVE: Cáncer de próstata, resonancia magnética bi-paramétrica (bp-MRI), segmentación de lesiones, mecanismo de atención de segundo orden

DESCRIPCIÓN: La detección de lesiones de cáncer de próstata clínicamente significativas (csPCa) a partir de imágenes de resonancia magnética biparamétrica (bp-MRI) ha surgido como una técnica de imagen no invasiva para mejorar la precisión diagnóstica. No obstante, el análisis de dichas imágenes sigue siendo altamente dependiente de la interpretación subjetiva del experto. Se han propuesto enfoques basados en aprendizaje profundo para la detección y segmentación de lesiones csPCa, pero estos permanecen limitados debido a su dependencia de conjuntos de datos extensamente anotados. Además, la alta variabilidad de las lesiones entre las diferentes zonas prostáticas plantea desafíos adicionales, incluso para radiólogos expertos. Este trabajo introduce un mecanismo de atención geométrica de segundo orden (SOGA, Second-Order Geometric Attention), que guía una red de segmentación, concretamente en sus conexiones de salto, para detectar y segmentar lesiones csPCa. La atención propuesta se modela en una variedad Riemanniana, aprendiendo a partir de representaciones simétricas definidas positivas (SPD). El mecanismo propuesto fue integrado en arquitecturas base tipo U-Net y nnU-Net, y validado en el conjunto de datos público PI-CAI, alcanzando una Precisión Promedio (AP) de 0.37 y un Área Bajo la Curva ROC (AUC-ROC) de 0.83, superando a la línea base y a los métodos con mecanismos de atención convencionales. Además, el enfoque fue evaluado en el conjunto de datos Prostate158 como cohorte de prueba independiente, obteniendo un AP de 0.37 y un AUC-ROC de 0.75, lo que confirma su capacidad de generalización robusta y sugiere el aprendizaje de representaciones discriminativas.

* Trabajo de investigación

** Facultad de Ingenierías Fisicomecánicas. Escuela de Ingeniería de Sistemas e Informática. Director: Fabio Martínez, PhD. Codirector: Juan Andrés Olmos Rojas, Msc.

ABSTRACT

TITLE: Segmentation of prostate cancer lesions in bi-parametric magnetic resonance imaging using second-order attention mechanisms. *

AUTHOR: Mateo Gerónimo Ortiz Cruzate **

KEYWORDS: Prostate cancer, bi-parametric resonance imaging (bp-MRI), lesion segmentation, second-order attention mechanism

DESCRIPTION: The detection of clinically significant prostate cancer lesions (csPCa) from bi-parametric magnetic resonance imaging (bp-MRI) has emerged as a noninvasive imaging technique for improving accurate diagnosis. Nevertheless, the analysis of such images remains highly dependent on the subjective expert interpretation. Deep learning approaches have been proposed for csPCa lesions detection and segmentation, but they remain limited due to their reliance on extensively annotated datasets. Moreover, the high lesion variability across prostate zones poses additional challenges, even for expert radiologists. This work introduces a second-order geometric attention (SOGA) mechanism that guides a dedicated segmentation network, through skip connections, to detect csPCa lesions. The proposed attention is modeled on the Riemannian manifold, learning from symmetric positive definitive (SPD) representations. The proposed mechanism was integrated into standard U-Net and nnU-Net backbones, and was validated on the publicly available PI-CAI dataset, achieving an Average Precision (AP) of 0.37 and an Area Under the ROC Curve (AUC-ROC) of 0.83, outperforming baseline networks and attention-based methods. Furthermore, the approach was evaluated on the Prostate158 dataset as an independent test cohort, achieving an AP of 0.37 and an AUC-ROC of 0.75, confirming robust generalization and suggesting discriminative learned representations.

* Research work

** Faculty of Physics-Mechanics Engineering. School of Systems Engineering and Informatics. Advisor: Fabio Martínez Carrillo, PhD. Co-advisors: Juan Andrés Olmos Rojas, Msc.

INTRODUCCIÓN

El cáncer de próstata (PCa) es el segundo cáncer más común y representa la quinta causa principal de muerte relacionada con cáncer en hombres. En 2022, se reportaron más de 1,5 millones de nuevos casos y más de 390.000 muertes a nivel mundial ¹. La detección de PCa se basa comúnmente en la prueba de antígeno prostático específico (PSA), pero presenta una baja especificidad, cercana al 36%, lo que conduce a numerosos falsos positivos y procedimientos innecesarios ². El examen rectal digital (DRE) también se utiliza, pero es invasivo, subjetivo y limitado en las regiones más allá de la pared rectal ².

Recientemente, la resonancia magnética multiparamétrica (mp-MRI) ha demostrado un apoyo significativo en el diagnóstico previo a la biopsia, al permitir la localización de lesiones de cáncer de próstata clínicamente significativas (csPCa) y resaltar tanto propiedades vasculares como morfológicas del tejido ³. No obstante, la mp-MRI implica tiempos de adquisición prolongados y el uso de agentes de contraste. Como alternativa, la resonancia magnética biparamétrica (bp-MRI) excluye la mejora dinámica con contraste (DCE) y ofrece una precisión diagnóstica comparable, proporcionando una solución práctica para el cribado a gran escala y el uso clínico rutinario ⁴. Sin embargo, en ambos casos, mp-MRI

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 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.

³ James THOMPSON et al. "The role of magnetic resonance imaging in the diagnosis and management of prostate cancer". In: *BJU international* 112 (2013), pp. 6–20.

⁴ Marco GATTI et al. "Prostate cancer detection with biparametric magnetic resonance imaging (bpMRI) by readers with different experience: performance and comparison with multiparametric (mpMRI)". In: *Abdominal Radiology* 44 (2019), pp. 1883–1893.

y bp-MRI, la interpretación y caracterización de las lesiones es subjetiva y variable, ya que depende del experto, lo que con frecuencia provoca confusión entre lesiones malignas e hiperplasia prostática^{5, 4, 3}.

En los últimos años, se han propuesto principalmente arquitecturas basadas en U-Net para la detección de csPCa en bp-MRI, debido a su capacidad para capturar características contextuales multiescala⁶. Otros trabajos han integrado módulos de atención que refinan las representaciones a nivel de canal y espacial durante la fusión de las características del codificador y el decodificador, aunque dependen de grandes conjuntos de datos anotados^{7, 8}. Cuando se entrenan con datos limitados o cohortes de un único centro, comúnmente los modelos basados en atención suelen mostrar un rendimiento y una capacidad de generalización reducidos, especialmente en conjuntos de datos externos⁹. En este trabajo de grado se propuso un módulo de atención geométrica de segundo orden (SOGA, por sus siglas en inglés) destinado a capturar representaciones profundas de mayor relevancia en bp-MRI, con el fin de mejorar la detección de lesiones csPCa. El mecanismo *SOGA* reduce los bancos de características en descriptores compactos de segundo orden, resumiendo las correlaciones entre características. Estos descriptores re-

⁵ Jonathan SHOAG and Christopher E BARBIERI. “Clinical variability and molecular heterogeneity in prostate cancer”. In: *Asian journal of andrology* 18.4 (2016), pp. 543–548.

⁶ Olaf RONNEBERGER; Philipp FISCHER, and Thomas BROX. “U-net: Convolutional networks for biomedical image segmentation”. In: *Medical image computing and computer-assisted intervention—MICCAI 2015: 18th international conference, Munich, Germany, October 5-9, 2015, proceedings, part III 18*. Springer. 2015, pp. 234–241.

⁷ 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.

⁸ Abdul Mueed HAFIZ; Shabir Ahmad PARAH, and Rouf UI Alam BHAT. “Attention mechanisms and deep learning for machine vision: A survey of the state of the art”. In: *arXiv preprint arXiv:2106.07550* (2021).

⁹ Audrey DURAN et al. “ProstAttention-Net: A deep attention model for prostate cancer segmentation by aggressiveness in MRI scans”. In: *Medical Image Analysis* 77 (2022), p. 102347.

sultantes son matrices simétricas definidas positivas (SPD) que coexisten en una variedad Riemanniana. En consecuencia, se consideraron capas de aprendizaje profundo que respetan su estructura para aprender descriptores geométricos SPD. Posteriormente, las representaciones SPD obtenidas se proyectan a un espacio euclidiano y se utilizan para refinar los bancos de características, permitiendo a la red aprender patrones más discriminativos a partir de la información de segundo orden. El módulo propuesto (*SOGA*) se integró en una arquitectura U-Net estándar y también se incorporó dentro del entorno nnU-Net. La inclusión de *SOGA* demostró la capacidad de aprender características más discriminativas y generalizables, mejorando el rendimiento en la detección y segmentación, destacando en la evaluación con cohortes externas.

1. FUNDAMENTOS Y TRABAJOS PREVIOS

1.1. CÁNCER DE PRÓSTATA Y SECUENCIAS MRI

El desarrollo del cáncer de próstata (*PCa*, por sus siglas en inglés) implica la proliferación de células dentro de la glándula prostática, dando lugar a la formación de tumores malignos ¹⁰. La aparición de este tipo de cáncer impacta aspectos fundamentales del paciente, como la salud reproductiva, el estado emocional y la calidad de vida ^{11, 12,13}. En sus etapas iniciales, el cáncer de próstata suele ser asintomático o incluir síntomas leves como molestias en la zona pélvica, dificultad para orinar o presencia de sangre en la orina ¹⁴. En fases avanzadas, la enfermedad puede provocar síntomas más graves, como dolor óseo por metástasis, pérdida de peso, fatiga, disfunción eréctil, entre otros ¹⁵. La biopsia es el método de referencia (*gold standard*) para el diagnóstico del cáncer de próstata. Este procedimiento consiste en la obtención de muestras de tejido prostático, que posteriormente son analizadas histopatológicamente para la confirmación de la sospecha de cáncer. Sin embargo, existen efectos adversos que pueden surgir durante su realización, como el

¹⁰ Amit R PATEL and Eric A KLEIN. “Risk factors for prostate cancer”. In: *Nature clinical practice Urology* 6.2 (2009), pp. 87–95.

¹¹ Lee Anne WALMSLEY. *The psychological well-being of men diagnosed with prostate cancer*. University of Kentucky, 2015.

¹² AnnMarie GROARKE et al. “Quality of life and adjustment in men with prostate cancer: Interplay of stress, threat and resilience”. In: *Plos one* 15.9 (2020), e0239469.

¹³ Rodolfo Alberto REY et al. “Male hypogonadism: an extended classification based on a developmental, endocrine physiology-based approach”. In: *Andrology* 1.1 (2013), pp. 3–16.

¹⁴ Elisabeth JS KUNKEL et al. “Biopsychosocial aspects of prostate cancer”. In: *Psychosomatics* 41.2 (2000), pp. 85–94.

¹⁵ Joseph A SMITH JR; Mark S SOLOWAY, and Marvin J YOUNG. “Complications of advanced prostate cancer”. In: *Urology* 54.6 (1999), pp. 8–14.

sangrado, infecciones, dolor y posibles dificultades para orinar después de la intervención¹⁶. Además, este procedimiento, de acuerdo con la rutina clínica, debe darse después de un periodo de tamizaje y evaluación con otros métodos, retrasando la planificación y tratamiento de la enfermedad. Dentro de los métodos de diagnóstico temprano del riesgo de cáncer de próstata se incluyen la medición del antígeno prostático específico (PSA) y el examen digital rectal (DRE). Aunque el uso de estas técnicas ha demostrado reducir las tasas de mortalidad, persiste un número significativo de falsos positivos².

Recientemente, la resonancia magnética (*MRI*, por sus siglas en inglés) ha emergido como una herramienta no invasiva clave para la visualización y estratificación del cáncer en etapas previas a la biopsia. Esta técnica facilita la caracterización del tejido, la localización de neoplasias y la estimación de su tamaño y grado de agresividad^{17, 18}. La imagenología biparamétrica (bp-MRI), compuesta exclusivamente por las secuencias T2W, DWI y mapas ADC, ha sido propuesta como una alternativa a la mp-MRI debido a su configuración simplificada que excluye la adquisición de imágenes dinámicas con contraste (DCE)^{19, 20}. El uso de bp-MRI ha demostrado disminuir el tiempo y los costos de captura, así como también evita el uso de gadolinio, un agente de contraste que puede generar

¹⁶ Stacy LOEB et al. "Systematic review of complications of prostate biopsy". In: *European urology* 64.6 (2013), pp. 876–892.

¹⁷ Richard J REBELLO et al. "Prostate cancer". In: *Nature Reviews Disease Primers* 7.1 (2021), p. 9. DOI: [10.1038/s41572-020-00243-0](https://doi.org/10.1038/s41572-020-00243-0).

¹⁸ Amita SHUKLA-DAVE and Hedvig HRICAK. "Role of MRI in prostate cancer detection". In: *NMR in Biomedicine* 27.1 (2014), pp. 16–24.

¹⁹ 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.

²⁰ Michele SCIALPI et al. "Biparametric MRI of the prostate". In: *Turkish journal of urology* 43.4 (2017), p. 401.

efectos adversos en los pacientes ^{20, 21}. Para la lectura e interpretación de hallazgos relacionados con lesiones de cáncer de próstata en estudios imagenológicos, los especialistas emplean el protocolo *Prostate Imaging Reporting and Data System* (PI-RADS) versión 2.1 ²². Este sistema proporciona un marco que estandariza la interpretación de los hallazgos por parte de los expertos, asignando puntuaciones que indican el grado de sospecha de lesión csPCa en regiones de interés. Específicamente, PI-RADS utiliza una escala en tres niveles de sospecha: baja (categorías 1 o 2), intermedia (categoría 3) y alta (categorías 4 o 5) ²³. Para estimar esta puntuación se consideran diversos parámetros morfológicos y funcionales. En la secuencia T2 ponderada (T2W) se examinan las características morfológicas de las zonas periférica y central de la próstata. En la secuencia ponderada por difusión (DWI) se analiza la restricción de la difusión de moléculas de agua, lo cual permite identificar áreas sospechosas de malignidad. Esta evaluación se complementa con el mapa de coeficiente de difusión aparente (ADC), donde las lesiones csPCa suelen presentar valores reducidos. Estos parámetros combinados contribuyen a la puntuación PI-RADS, la cual facilita la clasificación de la probabilidad de malignidad de las lesiones prostáticas y orienta el diagnóstico. A continuación, se describirán en detalle cada una de estas secuencias radiológicas de los estudios bp-MRI, así como sus características para la detección de lesiones csPCa. Además, se explicará el rol que desempeñan las distintas zonas prostáticas en la interpretación de estas imágenes.

²¹ Helen XU et al. "Prostate cancer detection using residual networks". In: *International journal of computer assisted radiology and surgery* 14 (2019), pp. 1647–1650.

²² R SÁNCHEZ-ORO et al. "Prostate MRI: practical guidelines for interpreting and reporting according to PI-RADS version 2.1". In: *Radiología (English Edition)* 62.6 (2020), pp. 437–451.

²³ 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.

1.1.1. Secuencia T2 ponderada La imagen ponderada en T2 (T2W) es una técnica que permite visualizar la anatomía de la glándula prostática ²⁴. Su adquisición implica la aplicación de un campo magnético que interactúa con los protones de agua en los tejidos, midiendo el tiempo de relajación tras la interrupción del pulso de radiofrecuencia. Debido a su alta resolución espacial y de contraste, permite evaluar la morfología zonal prostática ²⁵ y detectar lesiones sospechosas mediante el análisis de diferentes planos de visualización y focos hipointensos, los cuales sugieren hallazgos de lesión ^{26,24}. A pesar de su utilidad, las lesiones benignas relacionadas con otras enfermedades de la próstata, como la prostatitis o la hiperplasia prostática, pueden presentar patrones similares a los de las lesiones csPCa, dificultando su diferenciación ³. En la Figura 1 se ilustran tres casos con su respectiva visualización de la secuencia T2W. El paciente A presenta regiones con intensidades regulares y homogéneas en su glándula prostática. En contraste, el paciente B exhibe una morfología con intensidades irregulares que podrían interpretarse como una lesión csPCa, aunque el diagnóstico real corresponde a una hiperplasia prostática benigna (HPB).

Secuencia de difusión ponderada y mapas ADC

La imagen de difusión ponderada (DWI) se basa en el principio físico de la difusión, que describe el movimiento aleatorio de las moléculas de agua, conocido como movimiento

²⁴ Gillian MURPHY et al. "The expanding role of MRI in prostate cancer". In: *American Journal of Roentgenology* 201.6 (2013), pp. 1229–1238.

²⁵ Raquel 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.

browniano^{27, 28}. En condiciones normales, como en tejido prostático sano, este movimiento es libre y la señal en las imágenes DWI es baja. En contraste, cuando hay una alta densidad celular, como en tumores malignos, las moléculas encuentran barreras a su desplazamiento, lo que reduce la difusión. Esta restricción se traduce en un aumento de la señal en las imágenes DWI, generando una apariencia hiperintensa en las lesiones csPCa. De esta manera, la imagen DWI se ha convertido en una herramienta fundamental para identificar lesiones relacionadas con el cáncer de próstata²⁹. Recapitulando, las áreas con restricción de difusión aparecen como regiones hiperintensas, mientras que las hipointensas suelen asociarse a tejidos benignos³⁰.

Un parámetro clave para esta secuencia es el valor b (*b-value*), que puede ajustarse durante la adquisición para controlar la intensidad y duración de los gradientes aplicados³¹. Lo anterior permite generar los mapas de Coeficiente de Difusión Aparente (ADC), los cuales reflejan la movilidad molecular. De modo que un ADC bajo indica una difusión reducida, asociándose generalmente con tumores o progresiones cancerígenas, ya que el incremento en la densidad celular disminuye el espacio intercelular y extracelular³²,

²⁷ Olaf DIETRICH et al. "Technical aspects of MR diffusion imaging of the body". In: *European journal of radiology* 76.3 (2010), pp. 314–322.

²⁸ Martin H MAURER and Johannes T HEVERHAGEN. "Diffusion weighted imaging of the prostate—principles, application, and advances". In: *Translational andrology and urology* 6.3 (2017), p. 490.

²⁹ 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.

³⁰ Pilar Vallejo DESVIAT et al. "Difusión en resonancia magnética en el estudio de lesiones hepáticas". In: *Cirugía Española* 91.1 (2013), pp. 9–16.

³¹ I CAGLIC and T BARRETT. *Diffusion-weighted imaging (DWI) in lymph node staging for prostate cancer. Transl Androl Urol.* 2018; 7 (5): 814–23.

³² Tom J SYER et al. "The diagnostic accuracy of high b-value diffusion-and T 2-weighted imaging for the detection of prostate cancer: a meta-analysis". In: *Abdominal Radiology* 43 (2018), pp. 1787–1797.

^{28,33}. De esta manera, para estudio del PCa, los mapas ADC permiten soportar la discriminación entre tejidos malignos y sanos, donde las regiones con difusión restringida aparecen como áreas hipointensas ^{34, 30}. No obstante, una limitación importante es su baja resolución en comparación con otras secuencias, lo que puede dificultar una caracterización precisa de la malignidad por parte de los expertos ³⁵. En la Figura 1 se presentan también las secuencias DWI y ADC para los tres casos de ejemplo. En estas imágenes puede observarse que para el paciente C las lesiones csPCa se manifiestan como regiones hiperintensas en DWI, mientras que en su respectivo mapa de ADC se observan áreas hipointensas.

1.1.2. Zonas prostáticas La próstata se divide principalmente en dos zonas: la zona central (CZ) y la zona periférica (PZ). La zona periférica es la más grande, cubriendo alrededor del 70-80% del total del tamaño de la glándula prostática y se encuentra en la parte exterior de la próstata. Aproximadamente el 85% del cáncer de próstata diagnosticado se encuentra en esta zona, incluyendo las lesiones csPCa ³⁶. Por otro lado, la zona central tiene una menor prevalencia de patologías malignas, siendo más frecuente en ella la presencia de benignidades, como anormalidades asociadas a la hiperplasia prostática

³³ Michael A JACOBS et al. "Diffusion weighted imaging with adc mapping and spectroscopy in prostate cancer". In: *Medicine, Baltimore* 1500 (), p. 21205.

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

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.

³⁶ Grace L LU-YAO et al. "Outcomes of localized prostate cancer following conservative management". In: *Jama* 302.11 (2009), pp. 1202–1209.

benigna (HPB) ³⁷. En la Figura 1 se visualizan tres ejemplos en los que se observan delineaciones de la zona central y periférica sobre la secuencia T2W. El conocimiento sobre la localización de estas regiones aportan información anatómica clave que puede guiar al radiólogo a un análisis más informado, dado que la detección de lesiones csPCa suele ser de mayor facilidad en la zona periférica. En cambio, en la zona central, el tejido presenta una mayor variabilidad en la intensidad debido al desarrollo de tejidos benignos como la HPB, lo cual puede dificultar la diferenciación entre lesiones malignas y estructuras normales.

1.2. ARQUITECTURAS PARA LA SEGMENTACIÓN DE IMÁGENES MÉDICAS

1.2.1. Arquitectura U-Net Esta arquitectura autocodificadora fue propuesta para abordar tareas de segmentación sobre imágenes biomédicas, convirtiéndose en uno de los enfoques más utilizados⁶. Su estructura, codificador-decodificador, representado típicamente en forma de “U”, está compuesta por dos partes principales: un codificador (camino de contracción) y un decodificador (camino de expansión), conectados mediante una sección intermedia conocida como “cuello de botella”. Además, incorpora conexiones de salto (*skip connections*) entre el codificador y decodificador, lo que permiten preservar detalles espaciales clave en la segmentación final del modelo. El codificador se encarga de extraer características relevantes de la imagen de entrada a través de una serie de bloques convolucionales que reducen progresivamente la resolución espacial mientras aumentan el número de canales. De esta manera, se construye una representación abstracta y enriquecida a partir de información espacial. El decodificador, por su parte, toma esta representación resultante y la reconstruye a la resolución original mediante operaciones de sobremuestreo y convoluciones, generando como salida un mapa de segmentación. Esta

³⁷ Amin ALI et al. “Prostate zones and cancer: lost in transition?” In: *Nature Reviews Urology* 19.2 (2022), pp. 101–115.

■ Zona central ■ Zona periférica ■ Lesión csPCa

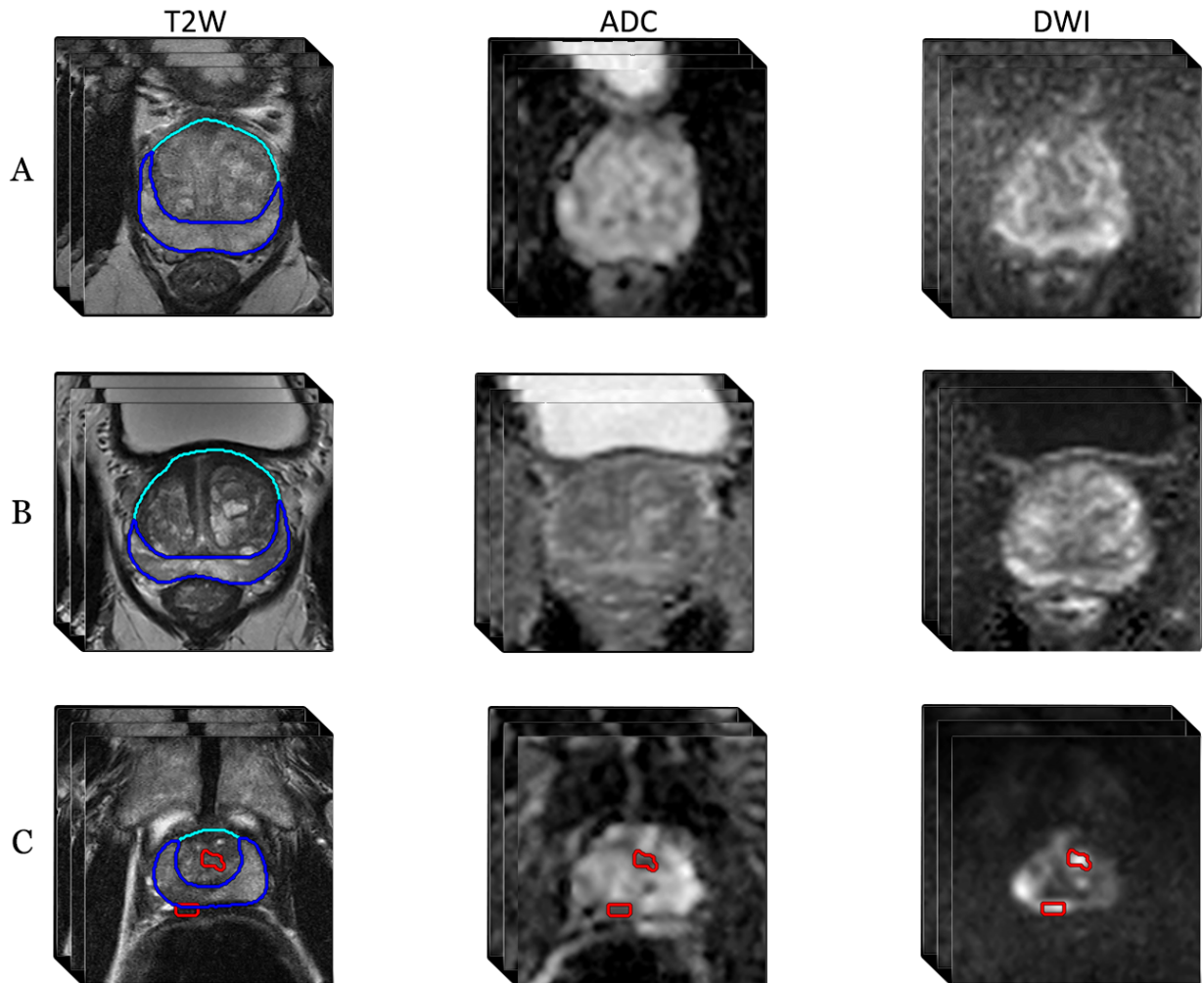


Figura 1. Ejemplos de hallazgos en estudios bp-MRI. La figura muestra imágenes de tres pacientes: a) Paciente control, b) Paciente con hiperplasia prostática benigna (HPB), y c) Paciente con dos lesiones csPCa. Las columnas corresponden a las secuencias T2W (izquierda), ADC (centro) y DWI (derecha). Se destacan también la zona central (cian), la zona periférica (azul) y, en el caso del paciente C, las lesiones csPCa (rojo).

reconstrucción se ve beneficiada por las conexiones de salto, que transfieren información espacial desde el codificador al decodificador, permitiendo al modelo recuperar detalles espaciales finos y facilitando el flujo de gradientes durante el entrenamiento, mejorando

así la capacidad de aprendizaje del modelo a diferentes niveles. De esta manera, dada una imagen $I \in \mathbb{R}^{W \times H}$, la U-Net produce un mapa de segmentación $a \in \mathbb{R}^{W \times H \times K}$ en la salida del último bloque del convolucional del decodificador, donde $a : \Omega \rightarrow \mathbb{R}^K$, siendo $\Omega \subset \mathbb{Z}^2$ las posiciones de los píxeles, y K el número de activaciones de salida de la red. Tradicionalmente cada activación representa un significado semántico sobre la tarea de segmentación. De esta manera, $a_k(\mathbf{x})$ denota la activación predicha en el canal de característica k en el píxel $\mathbf{x} \in \Omega$. Finalmente, para obtener un mapa de probabilidades y posteriormente, una máscara de segmentación, la U-Net propone aplicar una función softmax pixel-a-pixel como $p_k(\mathbf{x}) = \exp(a_k(\mathbf{x})) / \left(\sum_{k'=1}^K \exp(a_{k'}(\mathbf{x})) \right)$. De esta manera, cada pixel es comparado con una respectiva etiqueta semántica $\ell : \Omega \rightarrow \{0, \dots, N\}$. Para nuestro escenario de segmentación de lesiones csPCa, este problema se traduce en dos clases ($N = 2$), correspondientes a fondo y lesión. De esta manera, $p_k(\mathbf{x})$ se interpreta como la probabilidad de hallazgo de lesión a nivel de pixel. Comúnmente esta predicción es ajustada con una función de pérdida como la entropía cruzada ⁶. La Figura 2 ilustra el esquema codificador-decodificador de la U-Net, así como un ejemplo sobre la tarea de segmentación de lesiones csPCa sobre una secuencia MRI.

1.2.2. Método de segmentación nnU-Net El método nnU-Net (*no new U-Net*, por sus siglas en inglés) es un *framework* (marco) que permite realizar configuraciones en la arquitectura y el procesamiento, de forma semiautomática, para un problema específico de delineación en imágenes biomédicas. Este *framework* está diseñado para adaptarse automáticamente a las características específicas de cualquier conjunto de datos sin requerir ajustes manuales ³⁸. Su proceso comienza con la generación de un *dataset fingerprint*, en donde se extraen parámetros clave como tamaño y resolución de imágenes, modalidades, número de clases y estadísticas de las regiones de interés relacionadas

³⁸ Fabian ISENSEE et al. “nnU-Net: a self-configuring method for deep learning-based biomedical image segmentation”. In: *Nature methods* 18.2 (2021), pp. 203–211.

con la tarea de aprendizaje (distribución de la intensidad de la señal de entrada). Con esta información, nnU-Net genera un *pipeline fingerprint* a través de un conjunto de reglas heurísticas que definen aspectos de la arquitectura, como lo es el número de capas, tamaño de filtros convolucionales, número de características en cada nivel, entre otros. También considera el tipo de normalización, espaciado, tamaño de recorte a la zona de interés, y tamaño del lote de datos (*batch size*), entre otros. La arquitectura resultante deriva de una plantilla base tipo U-Net, sin agregar conexiones residuales, mecanismos de atención u otros módulos similares, priorizando en su lugar configuraciones fundamentadas como las mencionadas. Durante el entrenamiento, nnU-Net utiliza una pérdida de supervisión profunda, que consiste en medir la pérdida en distintos niveles de resolución mediante una combinación de la pérdida de entropía cruzada y una pérdida basada en la métrica *dice score*. También se considera un amplio rango de técnicas de aumento de datos como rotaciones, escalados, ruido Gaussiano, entre otros ³⁸. Durante la inferencia, se emplea una estrategia de ventana deslizante con filtro tipo gaussiano para reducir artefactos de borde y mejorar la robustez. Este enfoque, junto con estrategias de normalización específicas para cada tipo de imagen (imágenes de tomografía -CT-, o de resonancia magnética -MRI-, entre otras), permite a nnU-Net lograr un desempeño competitivo sin requerir intervención manual, lo cual lo convierte en un referente para tareas de segmentación médica automatizada ³⁸.

1.3. ESTRATEGIAS COMPUTACIONALES PARA LA SEGMENTACIÓN DE LESIONES CSPCA

La detección de lesiones csPCa en estudios bp-MRI representa un desafío fundamental para los sistemas de diagnóstico asistido por computadora. Esta tarea suele abordarse como un problema de segmentación, donde el objetivo es identificar las regiones del tejido prostático que presentan signos sospechosos de malignidad. En la literatura existen diversas estrategias computacionales para soportar esta tarea de segmentación, principal-

mente utilizando arquitecturas U-Net ⁶. Un ejemplo de estas es el trabajo de Chahal et al. quienes propusieron una arquitectura basada en U-Net y *Xception net*, donde a partir de la inclusión de conexiones residuales en el decodificador de la red lograron mejorar el desempeño del modelo ³⁹. Sin embargo, la segmentación de lesiones prostáticas implica desafíos particulares: las lesiones pueden ser pequeñas, de bajo contraste o estar ubicadas en zonas complejas anatómicamente, como la zona central, donde el entorno de la lesión puede confundirse con benignidades.

Para abordar estas limitaciones, se han propuesto mecanismos de atención, como el trabajo propuesto por Wei et al. donde se implementó un modelo basado en Attention U-Net, el cual aplica atención espacial refinando los bancos de características que se transfieren del codificador al decodificador, además, utiliza información de las zonas prostáticas en la entrada, guiando al modelo y mejorando la sensibilidad de detección de lesiones en la zona periférica ⁴⁰. Por otra parte, Yang, et al. utilizó una estrategia de atención por canal mediante bloques *Squeeze-and-Excitation* (SE) en las representaciones profundas del codificador de la red, las cuales mejoraron la extracción de características de parches de los estudios bp-MRI para la detección de PCa ⁴¹. Recientemente, Li et al propusieron un método basado en una arquitectura U-Net 3D, donde las máscaras anatómicas se incorporan como entrada adicional para guiar la segmentación de lesiones. Además, integraron bloques de atención tipo SE en el extractor de características y diseñaron una función de

³⁹ Ekam Singh CHAHAL et al. “Unet based xception model for prostate cancer segmentation from MRI images”. In: *Multimedia Tools and Applications* 81.26 (2022), pp. 37333–37349.

⁴⁰ Chao WEI et al. “Enhancing prostate cancer segmentation in bpMRI: Integrating zonal awareness into attention-guided U-Net”. In: *Digital Health* 11 (2025), p. 20552076251314546.

⁴¹ Chunguang YANG et al. “Deep learning model for the detection of prostate cancer and classification of clinically significant disease using multiparametric MRI in comparison to PI-RADs score”. In: *Urologic Oncology: Seminars and Original Investigations*. Vol. 42. 5. Elsevier. 2024, 158–e17.

pérdida que incorpora información sobre la localización zonal de las lesiones ⁴². Por otra parte, Duran et al. propusieron *ProstAttention-Net*, una red con dos decodificadores diseñada para segmentar conjuntamente la glándula prostática y las lesiones. Para ello, incorpora un mecanismo de atención espacial multiplicativa que conecta la representación final del decodificador encargado de segmentar la glándula prostática con el decodificador de las lesiones, mejorando así la segmentación de estas últimas y aprovechando la información anatómica de manera más explícita en el mecanismo de atención ⁹.

Recientemente, han surgido enfoques de diferentes trabajos que utilizan la nnU-Net. Por ejemplo, Debs et al. evaluaron el desempeño de este método, entrenando una 3D nnU-Net y corroborando su efectividad en la segmentación de lesiones csPCa en comparación con una red base U-Net ⁴³. De manera complementaria, Karagoz et al. incorporaron máscaras probabilísticas generadas automáticamente durante el entrenamiento de la nnU-Net, demostrando que la inclusión de las zonas prostáticas en este tipo de esquemas nnU-Net también permite guiar el modelo hacia una segmentación más apropiada ⁴⁴. Por otra parte, se han propuesto estrategias basadas en arquitecturas *transformers*, como el modelo CSwin-UNet propuesto por Li, et al., esta arquitectura considera un pre-entrenamiento auto-supervisado sobre un gran conjunto de bpMRI no anotadas, mejorando el desempeño sobre conjuntos de datos externos ⁴⁵.

Estas estrategias existentes resultan novedosas computacionalmente y han demostrado

⁴² Wenhao LI et al. "Adaptive window adjustment with boundary DoU loss for cascade segmentation of anatomy and lesions in prostate cancer using bpMRI". In: *Neural Networks* 181 (2025), p. 106831.

⁴³ Noëlie DEBS et al. "Evaluation of a deep learning prostate cancer detection system on biparametric MRI against radiological reading". In: *European Radiology* (2024), pp. 1–10.

⁴⁴ Ahmet KARAGOZ et al. "Anatomically guided self-adapting deep neural network for clinically significant prostate cancer detection on bi-parametric MRI: a multi-center study". In: *Insights into Imaging* 14.1 (2023), p. 110.

⁴⁵ Yuheng LI et al. "Cross-shaped windows transformer with self-supervised pretraining for clinically significant prostate cancer detection in bi-parametric MRI". In: *Medical Physics* 52.2 (2025), pp. 993–1004.

un desempeño sobresaliente para la segmentación de lesiones csPCa. Sin embargo, estos esquemas de tipo *transformer*, o en general arquitecturas de aprendizaje profundo con gran número de módulos a ajustar requieren necesariamente de una alta cantidad de datos, lo que en la práctica aún resulta limitado.

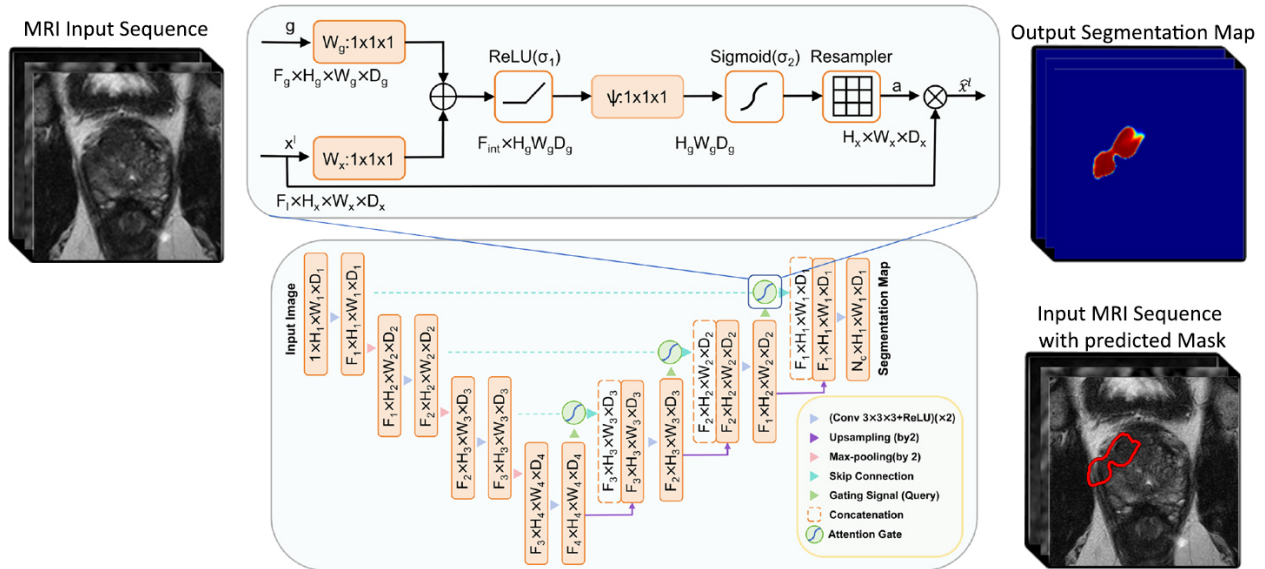


Figura 2. Ejemplo arquitectura U-Net para la segmentación de lesiones csPCa en bp-MRI. En la parte inferior se presenta la arquitectura U-Net, mientras en la parte superior se ilustra el mecanismo de atención introducido en *Attention U-Net*. A partir de una secuencia MRI (parte izquierda), se pretende que esta red produzca como salida un mapa de segmentación (parte derecha). Figura adaptada de *Wei et. al.*⁴⁶.

2. PROBLEMA DE INVESTIGACIÓN

El cáncer de próstata es el segundo con mayor incidencia a nivel mundial y la quinta causa de muerte por cáncer en hombres, con 1.5 millones de casos y más de 390,000 fallecimientos registrados en 2022 ¹. Tradicionalmente, el diagnóstico se basa en la medición del antígeno prostático específico (PSA) y el examen digital del recto (DRE), siendo este último un procedimiento altamente invasivo con limitada especificidad. La resonancia magnética biparamétrica (bp-MRI) ha mejorando la detección de lesiones clínicamente significativas de cáncer de próstata (csPCa), sin embargo, su análisis e interpretación es subjetivo. Recientemente, las herramientas computacionales han demostrado capacidad para soportar el diagnóstico de lesiones csPCa, pero las representaciones generadas ignoran el aprendizaje de relaciones entre características a lo largo de la arquitectura, lo que puede incurrir en un requerimiento excesivo de datos de entrenamiento. En la práctica clínica, esto se traduce en la necesidad de mas anotaciones, las cuales son limitadas debido al tiempo y costo de captura. Como alternativa, la integración de descriptores que exploten relaciones entre características, entre la representación visual y la de segmentación, podría mejorar la distinción entre tejidos sanos y malignos. Lo anterior nos conduce a formular la siguiente pregunta de investigación

Pregunta de investigación ¿Cómo integrar descriptores de segundo orden mediante un módulo de atención entre representaciones profundas biparamétricas para la segmentación de lesiones clínicamente significativas de cáncer de próstata?

3. OBJETIVOS

3.1. OBJETIVO GENERAL

Desarrollar un mecanismo de atención que integre descriptores de segundo orden para soportar la segmentación de lesiones clínicamente significativas de cáncer de próstata en secuencias de resonancia magnética biparamétrica.

3.2. OBJETIVOS ESPECÍFICOS

- Acondicionar un conjunto de secuencias bp-MRI que contenga delineaciones manuales de lesiones csPCa dadas por expertos radiólogos.
- Adaptar una arquitectura de aprendizaje profundo para la segmentación de lesiones csPCa en secuencias bp-MRI.
- Implementar un módulo de atención de segundo orden que capture relaciones entre representaciones profundas bp-MRI para soportar la segmentación de lesiones csPCa.
- Evaluar el desempeño del método propuesto para la segmentación y detección de lesiones csPCa.

4. MÉTODO PROPUESTO

En este trabajo se propuso un módulo de atención geométrica de segundo orden (*SOGA*, por sus siglas en inglés), el cual se integra en una representación codificador-decodificador para realizar la segmentación automática de lesiones de próstata. Específicamente, el módulo propuesto se incorpora entre las conexiones de salto (*skip connections*) entre el codificador y el decodificador, capturando correlaciones a nivel de características. El enfoque propuesto primero comprime mapas de características de alta dimensión en matrices simétricas definidas positivas (SPD), codificando de manera efectiva las correlaciones entre canales. Luego, el método aprovecha un módulo de procesamiento Riemanniano para preservar la geometría intrínseca de la variedad SPD durante todo el proceso de aprendizaje. Luego, los embebidos SPD se mapean a un espacio Euclidiano y se aplican proyecciones lineales para calcular pesos por canal que recalibrarán la información de las características que se transfieren a las capas subsiguientes en el decodificador. La Figura 3 ilustra en detalle el módulo *SOGA* propuesto, así como también la inclusión en la arquitectura dedicada a la segmentación de lesiones csPCa.

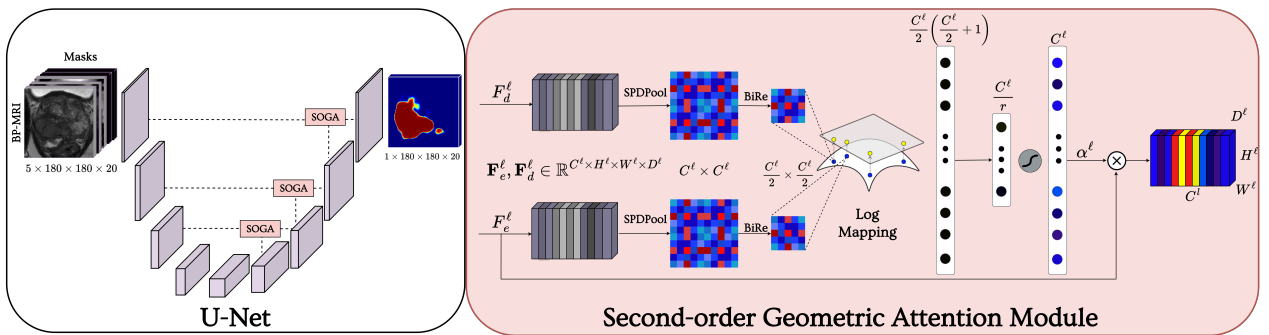


Figura 3. Figura del módulo de atención geométrica de segundo orden (*SOGA*) propuesto, integrado en una red U-Net.

4.1. REPRESENTACIÓN BASADA EN U-NET

Este trabajo considera arquitecturas actuales de segmentación *end-to-end* para la delimitación automática de lesiones csPCa en imágenes bp-MRI. Entre ellas, la U-Net sigue siendo el modelo más ampliamente adoptado para la segmentación de imágenes biomédicas debido a su capacidad para capturar contexto visual multiescala mediante conexiones de salto entre el codificador y el decodificador ⁶. Sin embargo, cabe destacar, que para la metodología propuesta, el módulo geométrico puede ser adaptado a cualquier metodología codificador-decodificador.

En resumen, la arquitectura U-Net consiste en un codificador que, en cada nivel ℓ , calcula progresivamente bloques de características denotados por $\mathbf{F}_e^\ell \in \mathbb{R}^{C_\ell \times H_\ell \times W_\ell \times D_\ell}$, donde C_ℓ es el número de canales de características en el nivel ℓ , $H_\ell \times W_\ell$ son las dimensiones espaciales y D_ℓ es la profundidad. Estas capas continúan hasta converger en una representación embebida (*bottleneck*), a partir de la cual el decodificador reconstruye la delimitación de la lesión mediante un *sobremuestreo* progresivo a través de bloques de características \mathbf{F}_d^ℓ . En cada nivel ℓ , las conexiones de salto transfieren información del codificador para complementar las representaciones contextuales del decodificador, mejorando así el rendimiento de la segmentación.

Recientemente, las conexiones de salto se han mejorado con la incorporación de mecanismos de atención, que acentúan la localización de estructuras relevantes mientras suprimen progresivamente el fondo irrelevante ^{47,48}. Este proceso implica recalibrar las activaciones de las características en alineación con la tarea objetivo. Con este fin, los mapas de características de entrada del codificador en cada nivel \mathbf{F}_e^ℓ se refinan mediante conex-

⁴⁷ Ozan OKTAY et al. "Attention u-net: Learning where to look for the pancreas". In: *arXiv preprint arXiv:1804.03999* (2018).

⁴⁸ Sanghyun WOO et al. "Cbam: Convolutional block attention module". In: *Proceedings of the European conference on computer vision (ECCV)*. 2018, pp. 3–19.

iones de salto basadas en atención, para luego incorporarse en la ruta del decodificador. La contribución de este trabajo radica en la propuesta de un mecanismo de atención geométrico de segundo orden en las conexiones de salto, que aprovecha las similitudes entre características para aprender representaciones de alta relevancia y capacidad discriminativa, mejorando así el desempeño de la red de segmentación base U-Net.

4.2. MECANISMOS DE ATENCIÓN Y APRENDIZAJE DE DESCRIPTORES DE SEGUNDO ORDEN

Los mecanismos de atención constituyen estrategias que permiten a las arquitecturas de aprendizaje automático identificar y enfocarse en la información más relevante de entrada ⁴⁹. En tareas de segmentación, este enfoque resulta crucial debido a la complejidad estructural de los objetos de interés —particularmente en imágenes biomédicas— y a la frecuente presencia de lesiones pequeñas o difusas. La atención puede aplicarse a nivel espacial, resaltando regiones anatómicas relevantes y suprimiendo ruido contextual, o a nivel de canales, ponderando la importancia relativa de cada mapa de características.

Un ejemplo representativo de atención espacial es *Attention U-Net* ⁴⁷, donde módulos en las conexiones de salto guían la transferencia de información desde el codificador hacia el decodificador. En este esquema, las representaciones del codificador x^ℓ son moduladas por coeficientes de atención $\alpha_i^\ell \in [0, 1]$, calculados en función de la información contextual proveniente del decodificador g^ℓ . De esta manera, se filtran las características transmitidas, reforzando regiones salientes y suprimiendo información irrelevante, lo que ha demostrado mejorar la localización de lesiones en contextos médicos ⁵⁰.

Por otro lado, mecanismos de atención a nivel de canal, como los bloques *Squeeze-and-*

⁴⁹ Ashish VASWANI et al. “Attention is all you need”. In: *Advances in neural information processing systems* 30 (2017).

⁵⁰ Hoel KERVADEC et al. “Beyond pixel-wise supervision for segmentation: A few global shape descriptors might be surprisingly good!” In: *Medical Imaging with Deep Learning*. PMLR. 2021, pp. 354–368.

Excitation (SE) ⁵¹, emplean operaciones de *global pooling* para condensar descriptores de cada canal en valores únicos que posteriormente se reescalan. No obstante, este enfoque basado en medidas de primer orden (media o máximo) puede provocar pérdida de información y una alta dependencia de la distribución de los mapas de características. Para mitigar esta limitación, en este trabajo se propuso un descriptor de segundo orden que permite modelar relaciones entre pares de características y capturar patrones más discriminativos incluso en conjuntos de datos reducidos ⁵². Formalmente, a partir de un banco de características $\mathbf{F} \in \mathbb{R}^{C \times H \times W}$, cada mapa $\mathbf{F}^{(i)}$ se vectoriza y organiza en $\mathbf{X}_{k-1} \in \mathbb{R}^{(H \cdot W) \times C}$. El descriptor de segundo orden se calcula como

$$\mathbf{X}_k = \frac{1}{W \times H} \mathbf{X}_{k-1}^\top \mathbf{X}_{k-1},$$

de esta manera, la entrada $\mathbf{X}_k(i, j)$ representa la correlación entre los mapas de características i y j . El resultado (\mathbf{X}_k) es una matriz simétrica definida positiva (SPD), la cual pertenece a una variedad Riemanniana ⁵³. Estos descriptores compactan de manera robusta información de alta dimensionalidad, preservando relaciones no lineales y permitiendo el uso de herramientas de aprendizaje geométrico profundo que respetan la estructura de este espacio ⁵⁴.

Para aprender nuevas representaciones geométricas a partir de este descriptor SPD, es

⁵¹ Jie HU; Li SHEN, and Gang SUN. “Squeeze-and-excitation networks”. In: *Proceedings of the IEEE conference on computer vision and pattern recognition*. 2018, pp. 7132–7141.

⁵² 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.

⁵³ William M BOOTHBY. *An introduction to differentiable manifolds and Riemannian geometry, Revised*. Vol. 120. Gulf Professional Publishing, 2003.

⁵⁴ Zhiwu HUANG and Luc VAN GOOL. “A riemannian network for spd matrix learning”. In: *Proceedings of the AAAI conference on artificial intelligence*. Vol. 31. 1. 2017.

necesario emplear operaciones conscientes con la geometría Riemanniana para preservar su estructura y permitir el aprendizaje de representaciones más significativas a partir de \mathbf{X}_0 ⁵⁴. Para el aprendizaje geométrico de nuevas matrices SPD (más compactas) se realiza una codificación en dos etapas, como se describe a continuación:

- *Capa BiMap*. Esta capa consiste en una proyección bilineal, definida como: $\mathbf{X}_k = f_{\text{BiMap}}(\mathbf{X}_{k-1}) = \mathbf{W}_k \mathbf{X}_{k-1} \mathbf{W}_k^\top$, donde $\mathbf{X}_{k-1} \in \mathcal{S}_{++}^{d_{k-1}}$ es la matriz SPD de la capa anterior, con dimensión $d_{k-1} \times d_{k-1}$, y $\mathbf{W}_k \in \mathbb{R}_*^{d_k \times d_{k-1}}$ es la matriz de transformación que genera la nueva matriz SPD $\mathbf{X}_k \in \mathcal{S}_{++}^{d_k}$. Al igual que en redes convencionales, las dimensiones de las matrices SPD se reducen progresivamente ($d_k < d_{k-1}$) mediante este mapeo bilineal.
- *Capa de rectificación ReEig*. Las matrices \mathbf{X}_k resultantes pueden tener autovalores cercanos a cero debido a procedimientos de optimización y numéricos durante el aprendizaje. Por ello, se considera una capa de rectificación de autovalores (*ReEig*) ⁵⁴, definida como: $\mathbf{X}_k = f_{\text{ReEig}}(\mathbf{X}_{k-1}) = \mathbf{U}_{k-1} \max(\epsilon \mathbf{I}, \Sigma_{k-1}) \mathbf{U}_{k-1}^\top$, donde \mathbf{U}_{k-1} y Σ_{k-1} corresponden a autovectores y autovalores respectivamente, con $\mathbf{X}_{k-1} = \mathbf{U}_{k-1} \Sigma_{k-1} \mathbf{U}_{k-1}^\top$. Aquí, $\epsilon \in \mathbb{R}$ es un umbral no negativo para la rectificación. Esta operación evita autovalores no positivos y preserva la estructura SPD ⁵⁴.

La concatenación de una capa *BiMap* y una *ReEig* se denomina bloque *BiRe*. Tras una secuencia de bloques *BiRe*, los descriptores Riemannianos se mapean de nuevo a un espacio euclidiano mediante el mapa logarítmico Riemanniano (*LogEig*): $\mathbf{X}_k = f_{\text{Log}}(\mathbf{X}_{k-1}) = \mathbf{U}_{k-1} \log(\Sigma_{k-1}) \mathbf{U}_{k-1}^\top$. La matriz \mathbf{X}_k resultante constituye un descriptor euclidiano de menor dimensión de relaciones semánticas entre características. Dado que \mathbf{X}_k es simétrica, se utiliza su parte triangular superior para el procesamiento posterior.

4.3. MÓDULO DE ATENCIÓN GEOMÉTRICA DE SEGUNDO ORDEN (SOGA)

Típicamente, los mecanismos de atención en redes tipo U-Net se implementan para guiar las características del codificador (\mathbf{F}_e^ℓ), usando características del decodificador (*Gating Signal*) (\mathbf{F}_d^ℓ) para inyectar contexto semántico y ayudar a la red a centrarse en regiones relevantes para la tarea de segmentación. Específicamente, la salida de un mecanismo de atención es $\hat{\mathbf{F}}^\ell = \mathbf{F}_e^\ell \odot \alpha_\ell$, donde α_ℓ es un vector aprendible con coeficientes de atención que refinan las activaciones del codificador. Estos coeficientes se calculan generalmente a partir de los tensores de características del codificador y decodificador, \mathbf{F}_e^ℓ y \mathbf{F}_d^ℓ , mediante operaciones de fusión como concatenación, multiplicación o suma. Como mencionamos en la sección anterior, este enfoque basado en descriptores de primer orden puede provocar pérdida de información y una alta dependencia de la distribución de los mapas de características.

En este trabajo, proponemos incluir información de segundo orden en todo el proceso de atención. Para ello, en cada nivel ℓ calculamos primero un descriptor SPD para los bancos de características del codificador y del decodificador como $\mathbf{X}_* = f_{SPDPool}(\mathbf{F}_*^\ell)$, con $* \in \{e, d\}$ denotando el codificador y decodificador respectivamente. Estos descriptores, de dimensión $C_\ell \times C_\ell$, capturan información relevante de los bancos de características, codificando patrones visuales semánticos clave propagados a través de las conexiones de salto y relacionados con el PCa. A partir de los descriptores SPD \mathbf{X}_e y \mathbf{X}_d , se obtienen representaciones geométricas de menor dimensión mediante dos redes SPD Riemannianas independientes φ_e y φ_d . Cada φ_* consiste en una bloque *BiRe* y una capa *LogEig* que produce una matriz simétrica de dimensión $\frac{C_\ell}{2} \times \frac{C_\ell}{2}$, y una capa de salida final que extrae la parte triangular superior. Las representaciones resultantes del codificador y decodificador se concatenan para formar una única representación geométrica embebida:

$$\mathbf{e}_\ell = [\varphi_e(f_{SPDPool}(\mathbf{F}_e^\ell)), \varphi_d(f_{SPDPool}(\mathbf{F}_d^\ell))] . \quad (1)$$

Estos descriptores fusionados producen un vector de dimensión $C_\ell(\frac{C_\ell}{2} + 1)$, que agrega información discriminativa del codificador y decodificador a través de descriptores SPD y procesamiento Riemanniano. Para obtener los coeficientes de atención, la representación \mathbf{e}_ℓ pasa por un mapeo no lineal compuesto por dos capas totalmente conectadas:

$$\alpha_\ell = \delta(\mathbf{W}_\ell^2 (W_\ell^1 \mathbf{e}_\ell + \mathbf{b}_\ell^1) + \mathbf{b}_\ell^2). \quad (2)$$

Aquí, $\mathbf{W}_\ell^1 \in \mathbb{R}^{\frac{C_\ell}{r} \times (C_\ell \cdot (\frac{C_\ell}{2} + 1))}$ y $\mathbf{W}_\ell^2 \in \mathbb{R}^{C_\ell \times \frac{C_\ell}{r}}$ representan las transformaciones lineales, y $b_\ell^1 \in \mathbb{R}^{\frac{C_\ell}{r}}$, $b_\ell^2 \in \mathbb{R}^{C_\ell}$ son los términos de sesgo. El hiperparámetro r controla la tasa de reducción, buscando equilibrar la capacidad representativa y la eficiencia en parámetros. Seleccionamos $r = 4$, como se describe en *Hu et. al.*⁵¹. El resultado de este proceso es el vector de coeficientes de atención que recalibra las características $\hat{\mathbf{F}}^\ell = \mathbf{F}_e^\ell \odot \alpha_\ell$.

4.4. DISEÑO EXPERIMENTAL

4.4.1. Datos. Se incluyeron dos conjuntos de datos distintos en la validación del enfoque propuesto, con el fin de evaluar las propiedades de generalización del método propuesto. La Figura 4 ilustra el protocolo experimental utilizado para validar el enfoque mediante los dos conjuntos de datos considerados. A continuación, describimos brevemente ambos conjuntos de datos.

Conjunto de datos del desafío PI-CAI. Este conjunto de datos es el mayor referente público para la detección de lesiones csPCa, compuesto por 1,500 estudios bp-MRI públicos para entrenamiento y desarrollo. Los organizadores proporcionaron las delimitaciones de las zonas central y periférica de la próstata para cada estudio. De los 1500 casos, 1075 son benignos (no-csPCa) y 425 corresponden a csPCa confirmados por biopsia. Entre los casos de csPCa, 220 incluyen anotaciones de expertos de las lesiones, mientras que los 205 restantes ofrecen delimitaciones generadas por IA denominadas “pseudo-etiquetas”, utilizadas exclusivamente para el entrenamiento. Para abordar la variabilidad en las imá-

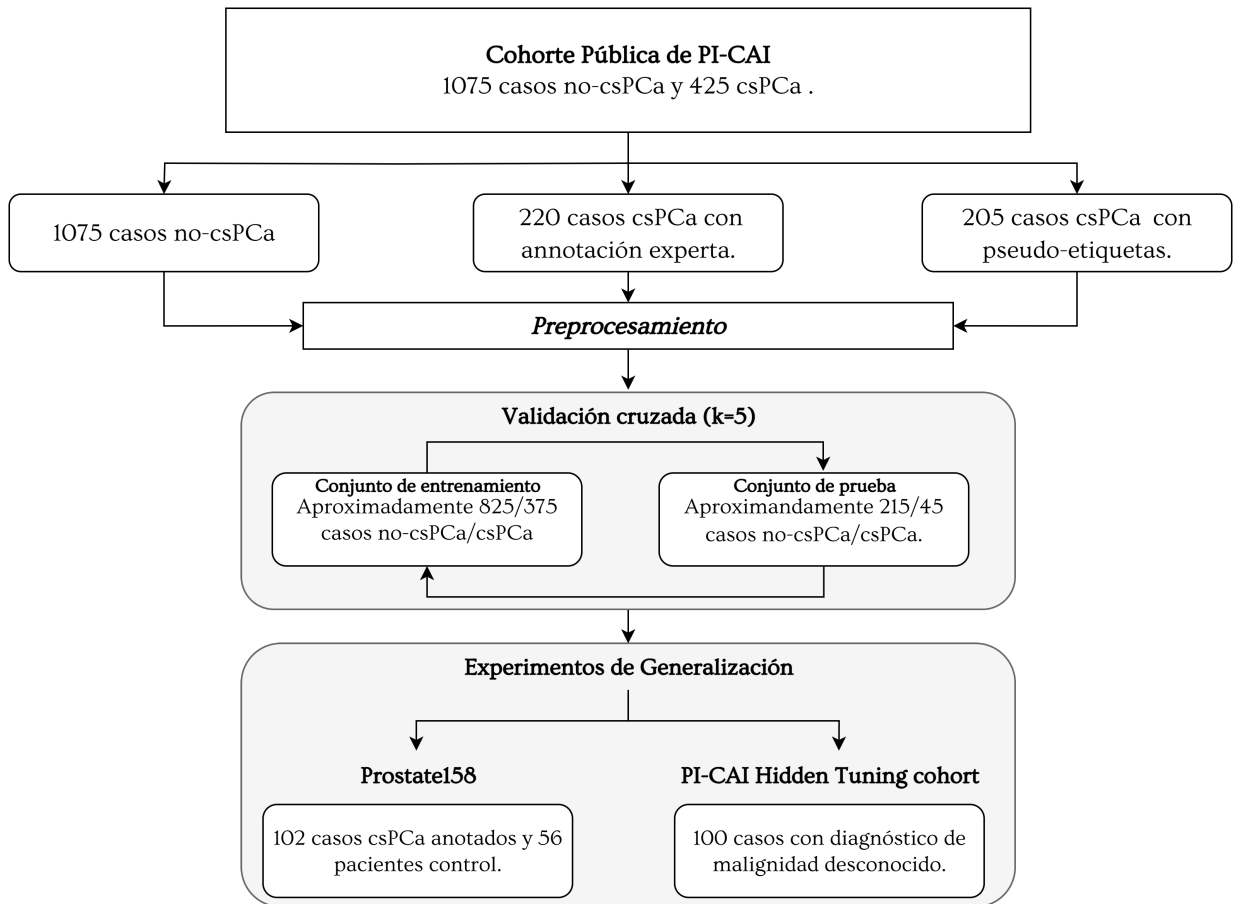


Figura 4. Diagrama de flujo de los conjuntos de datos utilizados para la validación. El entrenamiento y validación de los modelos se realizó en la cohorte pública PI-CAI. Adicionalmente, se evaluó la generalización de los modelos sobre una cohorte oculta de PI-CAI, y el conjunto de datos externo PROSTATE158.

genes, el conjunto incluye estudios bp-MRI de tres centros neerlandeses: Radboud University Medical Center (RUMC), Ziekenhuis Groep Twente (ZGT) y University Medical Center Groningen ⁵⁵.

El conjunto PI-CAI incluye una cohorte oculta adicional diseñada para el ajuste de mode-

⁵⁵ Anindo SAHA et al. "Artificial intelligence and radiologists in prostate cancer detection on MRI (PI-CAI): an international, paired, non-inferiority, confirmatory study". In: *The Lancet Oncology* 25.7 (2024), pp. 879–887.

los. Esta cohorte comprende 100 estudios bp-MRI. Los organizadores del desafío reservaron las delineaciones de las lesiones para este conjunto, y el rendimiento se evaluó de forma indirecta utilizando métricas reportadas por la plataforma de evaluación del desafío. Usamos esta cohorte para validar la capacidad de generalización de los modelos entrenados en la cohorte pública (véase la Figura 4).

Las etapas de preprocesamiento incluyeron el remuestreo de los estudios bp-MRI a una resolución de vóxel de $0.5\text{mm} \times 0.5\text{mm} \times 3.0\text{mm}$. Posteriormente, se realizó un recorte centrado para obtener 24 cortes axiales de 384×384 píxeles⁵⁵. Se obtuvo una región de interés volumétrica (vROI) de tamaño $20 \times 180 \times 180$ vóxeles alrededor del centroide volumétrico de la glándula prostática. Cada secuencia de entrada, incluyendo imágenes ponderadas en T2 (T2W), imágenes ponderadas por difusión (DWI) y mapas de coeficiente de difusión aparente (ADC), fueron normalizadas de forma independiente mediante normalización *z-score*.

PROSTATE158. Este conjunto de datos comprende estudios bp-MRI de 158 pacientes obtenidos en el Hospital Universitario Charité de Berlín, Alemania⁵⁶. Entre ellos, 102 casos fueron csPCa confirmados por biopsia o cirugía, mientras que los 56 restantes fueron controles sin cáncer. El conjunto está dividido en 139 casos para entrenamiento y 19 para prueba, todos anotados por dos radiólogos certificados. Cada caso incluye la delimitación experta de las zonas prostáticas (central y periférica). Se aplicaron los mismos pasos de preprocesamiento utilizados para la cohorte pública PI-CAI a esta cohorte externa. También utilizamos esta cohorte externa independiente para evaluar la generalización de nuestros modelos. Los casos de entrenamiento y prueba de PROSTATE158 se combinaron para formar un único conjunto de evaluación independiente. Una visión general de este flujo de trabajo se muestra en la Figura 4.

⁵⁶ Lisa C ADAMS et al. "Prostate158-An expert-annotated 3T MRI dataset and algorithm for prostate cancer detection". In: *Computers in Biology and Medicine* 148 (2022), p. 105817.

4.4.2. Configuración de las arquitecturas Seleccionamos una arquitectura 3D U-Net estándar para incluir el módulo de atención propuesto *SOGA*. La red consta de cinco niveles jerárquicos, con los siguientes números de canales: [32, 64, 128, 256, 512]. En cada nivel, esta arquitectura incluye un bloque convolucional con dos capas de convolución 3D (tamaño de filtro de $3 \times 3 \times 3$), seguido de normalización por lotes y una función de activación ReLU. La reducción de dimensionalidad se realizó mediante *MaxPooling* 3D (véase la Figura 3). Nos referimos a este modelo, sin bloques de atención, como *U-Net*. Para comparar la atención geométrica de segundo Orden (*SOGA*) propuesta, consideramos un módulo de atención base, utilizando *Global Average Pooling*. A este método lo denominamos *U-Net FOA* debido a que se basa en representaciones de primer orden de las características. Asimismo, el enfoque propuesto se comparó con una arquitectura que incluyó *SPD pooling* sin el posterior procesamiento geométrico, *i.e.*, sin la incorporación de bloques *BiRE* aprendibles. Este método se denomina *U-Net SOA* debido a que emplea una descripción de segundo orden de las características pero sin incluir el procesamiento geométrico.

En el método *SOGA* propuesto, se aplicó un umbral de rectificación de $\epsilon = 10^{-4}$ en las capas ReEig. Para los parámetros no-Riemannianos se utilizó RMSprop con una tasa de aprendizaje de 10^{-4} , mientras que los pesos BiMap fueron optimizados mediante descenso de gradiente en la variedad de Stiefel, con una tasa de aprendizaje de 10^{-1} ⁵⁴. Todos los modelos se entrenaron durante 200 épocas utilizando como función de pérdida la suma de la *Dice Loss* y la Entropía Cruzada Binaria. Además, se realizaron experimentos integrando estos métodos en el *framework nnU-Net* ³⁸. También, los modelos basados en transformers UNETR ⁵⁷ y Swin UNETR ⁵⁸ se incorporaron al *framework nnU-Net* para su

⁵⁷ Ali HATAMIZADEH et al. “Unetr: Transformers for 3d medical image segmentation”. In: *Proceedings of the IEEE/CVF winter conference on applications of computer vision*. 2022, pp. 574–584.

⁵⁸ Ali HATAMIZADEH et al. “Swin unetr: Swin transformers for semantic segmentation of brain tumors in mri images”. In: *International MICCAI brainlesion workshop*. Springer. 2021, pp. 272–284.

comparación con métodos de atención del estado del arte. A diferencia de los enfoques implementados únicamente sobre la arquitectura U-Net original, estos métodos implementados en *nnU-Net* aprovechan las capacidades automáticas de auto-configuración, como la adaptación dinámica de datos, ofreciendo un flujo de entrenamiento más robusto y generalizable ³⁸. El término *Baseline* hace referencia a la nnU-Net estándar, utilizada como punto de comparación.

Evaluación. Para evaluar el rendimiento de los métodos estudiados en la segmentación y detección de csPCa, se consideraron múltiples métricas que ofrecen una evaluación a nivel de paciente y de lesión. Antes de calcular las métricas, los mapas de detección y las puntuaciones de confianza (valores máximos de predicción a partir de dichos mapas) se obtuvieron utilizando el postprocesamiento recomendado por PI-CAI ⁵⁵. Se consideraron las siguientes métricas de evaluación:

- El área bajo la curva ROC (*AUC-ROC*), por sus siglas en inglés, definida como:

$$AUC-ROC = \int_0^1 TPR(FPR) dFPR,$$

donde *TPR* representa la tasa de verdaderos positivos y *FPR* la tasa de falsos positivos. Esta métrica se calcula a partir de las puntuaciones de confianza generadas por el modelo en los mapas de detección. Permite medir la capacidad del modelo para discriminar estudios bpMRI con presencia de lesión frente a aquellos sin lesión. Sus valores oscilan entre 0 y 1: un valor de 1 indica discriminación perfecta, 0.5 corresponde a un clasificador aleatorio y 0 representa una clasificación completamente incorrecta. En este caso, valores más altos son preferibles, pues reflejan un mejor desempeño del modelo en la tarea de detección.

- La precisión promedio (AP , por sus siglas en inglés), definida como:

$$AP = \sum_{n=1}^N (R_n - R_{n-1})P_n,$$

donde P_n y R_n corresponden a la precisión y la sensibilidad en el umbral n . Se utilizó una intersección sobre unión (IoU , por sus siglas en inglés) mínima de 0.1 para considerar una predicción como detección espacial positiva, siguiendo recomendaciones de la literatura ⁴⁴. El AP es una métrica ampliamente usada en tareas de detección, pues resume en un único valor el compromiso entre precisión y sensibilidad a lo largo de diferentes umbrales: la precisión disminuye con los falsos positivos, y la sensibilidad aumenta con la correcta detección de lesiones. Así, valores más altos de AP indican un mejor desempeño del modelo para detectar lesiones csPCa en estudios bpMRI.

- La Sensibilidad a 1 Falso Positivo ($SEN@1FP$), definida como:

$$SEN@1FP = \frac{TP}{TP + FN} \quad \text{con la restricción de } FP \leq 1 \text{ por paciente,}$$

donde TP son los verdaderos positivos y FN los falsos negativos. Esta métrica se calcula a nivel de paciente y mide la sensibilidad del modelo bajo la condición de permitir como máximo un falso positivo por caso. Es particularmente relevante en escenarios médicos, ya que refleja directamente el impacto de las lesiones no detectadas: valores más bajos indican un mayor número de falsos negativos, mientras que valores más altos corresponden a un mejor desempeño del modelo en la detección de lesiones clínicamente significativas.

- El coeficiente de similitud de Dice (DSC), definido como:

$$DSC = \frac{2|X \cap Y|}{|X| + |Y|},$$

donde X representa los vóxeles predichos como positivos y Y los vóxeles de la anotación de referencia. De esta manera, la métrica determina el grado de superposición entre una predicción y la anotación de referencia, siendo 1 un solapamiento perfecto y 0 uno totalmente erróneo. A diferencia de las métricas anteriores, que evalúan la capacidad de detección, el DSC se centra en medir la calidad de la segmentación de las lesiones, es decir, qué tan bien el modelo logra delinear las regiones sospechosas de lesiones csPCa en la bpMRI.

En los experimentos del desafío PI-CAI, seguimos la validación cruzada oficial de 5 particiones, utilizando las divisiones predefinidas proporcionadas por los organizadores ⁵⁹. Para la evaluación de los modelos sobre las cohortes PI-CAI Hidden Tuning y Prostate158, las predicciones se combinaron en una sola predicción promediando las salidas de *softmax* de cada una de las 5 las particiones. Para evaluar los resultados de manera estadísticamente significativa, se utilizó la prueba de *DeLong*, de esta manera se pudo comparar las áreas bajo la curva ROC de los diferentes modelos. En particular, se verificó si las diferencias en el AUC-ROC entre el método propuesto y las líneas base de comparación podían atribuirse al azar.

⁵⁹ Anindo SAHA et al. "The PI-CAI challenge: public training and development dataset". In: *Zenodo, Jun (2022)*.

5. EVALUACIÓN Y RESULTADOS

5.1. DETECCIÓN EN LA COHORTE PÚBLICA DE PI-CAI

Los resultados en la Cohorte Pública de PI-CAI se presentan en la Tabla 1. En primera instancia, se realizó una comparación con la incorporación de estos módulos en una arquitectura estándar U-net, considerándola como una línea base para observar como los diferentes módulos podrían mejorar la segmentación. De manera interesante, la integración del módulo SOA (*U-Net SOA*) alcanzó el mayor rendimiento global, mientras que el módulo propuesto SOGA (*U-Net SOGA*) mostró mejoras consistentes en todas las métricas de evaluación en comparación con la U-Net base y el módulo de primer orden (*U-Net FOA*). Siendo el mejor en segmentación, pasando de un DSC de 0.12 a 0.31 (+158%). La incorporación del módulo SOGA incrementó el rendimiento de la *U-Net* en AUC-ROC de 0.76 a 0.82 (+7.9%), lo que resultó estadísticamente significativo ($p < 0.05$), indicando una mayor capacidad del modelo para distinguir entre casos de csPCa y no-csPCa. Además, se observó una mejora considerable en la métrica Sen@1FP, que aumentó de 0.61 a 0.74 (+21.3%). Adicionalmente, el DSC se incrementó en 0.08 (+23.5%) y mostró una reducción en su desviación estándar.

Tabla 1. Desempeño en la cohorte pública PI-CAI con U-Net. Comparación entre el módulo propuesto SOGA (marcado con *) y modelos basados en atención en una configuración U-Net. Las mejores métricas están resaltadas en **negrilla**.

Modelo	AUC-ROC	AP	Sen@1FP	DSC
U-Net	0.76 ± 0.03	0.31 ± 0.05	0.61 ± 0.14	0.34 ± 0.09
U-Net FOA	0.70 ± 0.04	0.21 ± 0.08	0.49 ± 0.11	0.25 ± 0.08
U-Net SOA	0.85 ± 0.02	0.38 ± 0.05	0.73 ± 0.07	0.22 ± 0.26
U-Net SOGA*	0.82 ± 0.03	0.35 ± 0.08	0.74 ± 0.06	0.42 ± 0.04

En una segunda instancia, en la Tabla 2 se realizó una validación y comparación con arquitecturas del estado del arte. Aquí se implementó el framework nnU-Net estándar,

sin módulos de atención, conexiones residuales o alguna configuración adicional en la arquitectura, denotado como *Baseline*. Además, se decidió implementar los módulos de atención en el framework nn-UNet. Como era de esperar, el uso de la nnU-Net mejoró el rendimiento de los métodos basados en U-Net, demostrando la ventaja de este marco para ajustar los datos según su variabilidad y configurar los esquemas de entrenamiento para las tareas de detección y segmentación. Cuando el módulo *SOGA* se integra en la nnU-Net, se evidencia una mejora en el rendimiento en detección, con una puntuación AP de 0.37 ± 0.05 (+5.7%), y en segmentación, con una puntuación DSC de 0.52 ± 0.02 (+23.8%), y el AUC-ROC se mantuvo similar al *Baseline*, ($p = 0.695$). Por el contrario, los modelos basados en transformers (UNETR y Swin UNETR) obtuvieron un rendimiento global inferior al del propuesto *SOGA*, sin diferencias estadísticamente significativas ($p = 0.140$ y $p = 0.067$, respectivamente), y solo lograron mejorar ligeramente el rendimiento del enfoque nnU-Net base (baseline) en las métricas AP y DSC. Este comportamiento en las arquitecturas transformers puede estar asociado al número de parámetros de la arquitectura comparado con los datos disponibles para el entrenamiento, lo cual sigue siendo un desafío en el área clínica.

Tabla 2. Desempeño en la cohorte pública PI-CAI con nnU-Net. Comparación entre el módulo propuesto *SOGA* (marcado con *), nn-UNet (baseline) y modelos basados en atención en el framework nnU-Net. Las mejores métricas están resaltadas en **negrilla**.

Modelo	AUC-ROC	AP	Sen@1FP	DSC
Baseline	0.83 ± 0.04	0.33 ± 0.09	0.82 ± 0.04	0.47 ± 0.03
UNETR	0.82 ± 0.03	0.35 ± 0.08	0.82 ± 0.02	0.49 ± 0.03
Swin UNETR	0.80 ± 0.03	0.34 ± 0.07	0.83 ± 0.04	0.50 ± 0.03
FOA	0.83 ± 0.03	0.36 ± 0.70	0.81 ± 0.05	0.51 ± 0.04
SOA	0.82 ± 0.04	0.35 ± 0.10	0.83 ± 0.07	0.51 ± 0.02
SOGA*	0.83 ± 0.03	0.37 ± 0.05	0.82 ± 0.02	0.52 ± 0.02

5.2. EVALUACIÓN DE GENERALIZACIÓN

En lo que respecta a la validación sobre el conjunto de datos *PI-CAI Hidden Tuning cohort*, la plataforma organizadora del reto PI-CAI únicamente proporcionó las métricas AUC-ROC y AP. Para este experimento, se validaron los modelos entrenados mencionados anteriormente, es decir, nnUnet (Baseline), y aquellos que incorporan atención, ver Tabla 3. En este caso, las puntuaciones de AUC-ROC indican un rendimiento de detección similar entre todos los métodos. Sin embargo, la incorporación del módulo de atención propuesto *SOGA* produjo una mejora notable en la métrica AP, aumentando de 0.37 a 0.46 (+24.3%) en comparación con la nnUNet (Baseline). Esta diferencia en el rendimiento de detección es mayor que la observada en la cohorte pública de PI-CAI cuando *SOGA* fue incorporado en la nnU-Net, lo que indica que el módulo propuesto *SOGA* es robusto ante casos no vistos y su ventaja en estos escenarios es más notoria. Esta mejora sustancial resalta la capacidad de generalización del método en escenarios no vistos. A pesar de su buen rendimiento, el método propuesto *SOGA* fue superado por Swin UNETR, el cuál alcanzó un AP de 0.54.

Tabla 3. Desempeño en la cohorte PI-CAI Hidden Tuning. Comparación de la nnU-Net *baseline*, y nnU-Net con mecanismos de atención, incluyendo el propuesto *SOGA*. Las mejores métricas están resaltadas en **negrilla**.

Modelo	AUC-ROC	AP
Baseline	0.78	0.37
UNETR	0.75	0.47
Swin UNETR	0.78	0.54
FOA	0.76	0.40
SOA	0.71	0.44
SOGA*	0.77	0.46

Para evaluar más a fondo el rendimiento de generalización, realizamos evaluaciones adicionales en una cohorte externa e independiente, el conjunto de datos **Prostate158**. Los resultados se muestran en la Tabla 4. En este caso, los modelos basados en atención demostraron de forma consistente un rendimiento de generalización superior con respecto al

método nnUNet (Baseline), destacándose el módulo propuesto *SOGA* como el de mejor desempeño, seguido por los métodos UNETR y Swin UNETR. El método Swin UNETR tuvo una diferencia estadística en el AUC-ROC al compararlo con el método propuesto ($p < 0.05$). Sin embargo, la diferencia para la UNETR no fue significativa ($p = 0.140$). En comparación con la línea base nnU-Net, la inclusión de *SOGA* produjo una mejora de 0.13 puntos (+21.0%) en AUROC, lo que resultó estadísticamente significativo ($p < 0.05$), 0.22 (+146.6%) en AP, 0.29 (+126.1%) en Sen@1FP y 0.19 (+158.3%) en DSC. Estos resultados resaltan la capacidad del módulo *SOGA* para apoyar la detección y segmentación de lesiones csPCa, especialmente en casos provenientes de cohortes externas e independientes a las cohortes utilizadas para el entrenamiento. Esto puede explicarse por la capacidad del método propuesto para capturar de manera más efectiva patrones visuales relevantes para la identificación de lesiones csPCa, lo que permite un mejor rendimiento en conjuntos de datos externos, independientes y no vistos previamente. Cabe destacar que en esta cohorte existe un número considerable de estudios bpMRI sin lesión csPCa, a diferencia de la cohorte PI-CAI Hidden, lo que hace que los resultados obtenidos sobre este conjunto de datos sean particularmente relevantes. A partir de ello, la desmejora observada en el desempeño del modelo Swin UNETR sugiere una tendencia hacia una mayor tasa de falsos positivos.

Tabla 4. Desempeño en el conjunto de datos Prostate158. Comparación entre la nnU-Net (*baseline*) y de los métodos basados en atención incorporados en este framework. Las mejores métricas están resaltadas en **negrilla**.

Modelo	AUC-ROC	AP	Sen@1FP	DSC
Baseline	0.62	0.15	0.23	0.12
UNETR	0.73	0.29	0.38	0.23
Swin UNETR	0.70	0.27	0.38	0.23
FOA	0.67	0.18	0.31	0.18
SOA	0.72	0.29	0.43	0.24
SOGA*	0.75	0.37	0.52	0.31

Adicionalmente, calculamos el desempeño de los modelos en Prostate158 diferenciando

por tamaño de las lesiones. Para ello, las lesiones del conjunto de datos se estratificaron según su volumen en: *pequeñas* ($< 931 \text{ mm}^3$), *medianas* ($931\text{--}2337 \text{ mm}^3$) y *grandes* ($> 2337 \text{ mm}^3$), utilizando como umbrales los percentiles 33 y 66. Como se muestra en la Tabla 5, el módulo propuesto *SOGA* superó de manera consistente a la línea base nnU-Net y al resto de métodos basados en atención en todos los tamaños de lesión. De manera interesante, para las lesiones pequeñas, típicamente las más difíciles de detectar y segmentar, *SOGA* obtuvo el mejor rendimiento, seguido en orden por SOA, UNETR, Swin UNETR, FOA y nnU-Net base. En el grupo de lesiones medianas se observó un orden de rendimiento similar. Además, se reportaron mejores puntuaciones de detección y segmentación para todos los modelos en comparación con sus respectivos resultados en el grupo de lesiones pequeñas. En el caso de lesiones grandes, todos los métodos alcanzaron un rendimiento superior al observado en lesiones medianas y pequeñas. Aunque *SOGA* continuó mostrando un alto rendimiento global, fue superado en las métricas AP por Swin UNETR. Nótese que la ventaja del método propuesto *SOGA* fue mayor en las lesiones pequeñas y medianas.

Cabe destacar que la ventaja del método propuesto (*SOGA*) fue más evidente en las lesiones pequeñas y medianas, superando consistentemente al segundo mejor (SOA), que también emplea descriptores de segundo orden. En el caso de las lesiones pequeñas, *SOGA* alcanzó un DSC de 0.22 (+22%), un AP de 0.38 (+52%) y una SEN@1FP de 0.50 (+61%), mostrando una mejora sustancial en comparación con SOA. Para las lesiones medianas, los resultados también fueron superiores, con un DSC de 0.33 (+10%), un AP de 0.44 (+5%) y una SEN@1FP de 0.50 (+4%). Estos hallazgos evidencian no solo la capacidad de generalización del modelo propuesto, sino también su notable ventaja en la detección de lesiones pequeñas, un aspecto de gran utilidad en la práctica clínica, donde estas suelen pasar desapercibidas.

La Figura 5 presenta una comparación visual entre las anotaciones de expertos y las predicciones de los modelos para tres casos seleccionados aleatoriamente del conjunto

Tabla 5. Desempeño según diferentes tamaños de lesión en Prostate158. Comparación entre la nnU-Net *baseline* y los métodos basados en atención para la detección y segmentación de lesiones pequeñas, medianas y grandes. Las mejores métricas están resaltadas en **negrilla**.

Métrica	Baseline	UNETR	S. UNETR	FOA	SOA	SOGA*
<i>Pequeñas</i>						
DSC	0.10	0.14	0.11	0.11	0.18	0.22
AP	0.04	0.15	0.12	0.08	0.25	0.38
Sen@1FP	0.14	0.28	0.20	0.17	0.31	0.50
<i>Medianas</i>						
DSC	0.17	0.28	0.28	0.24	0.30	0.33
AP	0.26	0.38	0.36	0.33	0.42	0.44
Sen@1FP	0.28	0.41	0.41	0.37	0.48	0.50
<i>Grandes</i>						
DSC	0.16	0.26	0.31	0.24	0.29	0.40
AP	0.24	0.37	0.47	0.27	0.35	0.43
Sen@1FP	0.25	0.43	0.50	0.36	0.48	0.57

de datos Prostate158. En el caso a), se observa una lesión grande, delineada únicamente por Swin UNETR, SOA y el método propuesto SOGA, con una ligera ventaja en segmentación para SOGA de 3 puntos en el DSC. Además, existe una lesión de tamaño mediano que fue omitida por todos los modelos en todos los volúmenes 3D predichos. La segmentación no es visible en el corte actual ni en los adyacentes. Al observar con mayor detalle, se aprecia que en este corte la lesión aparece fragmentada en dos regiones separadas, lo cual pudo haber influido en que los modelos no la identificaran como una única estructura continua. El caso b) presenta una lesión de tamaño mediano detectada por todos los modelos. La nnU-Net (Baseline) solo capturó una porción de la lesión, a diferencia de los métodos basados en atención. De forma destacada, SOGA muestra una ventaja tanto en la detección de la lesión como en la mejor delineación de su morfología, obteniendo un DSC superior. El caso c) presenta una lesión pequeña en la zona periférica. Esta lesión fue detectada débilmente por los modelos nnU-Net y SOA, y omitida por completo por el modelo FOA, pero identificada exitosamente por los métodos UNETR, Swin

UNETR y *SOGA*. El método *SOGA* proporcionó una delineación más precisa y una detección más confiable, demostrando la ventaja de utilizar atención basada en descriptores de segundo orden frente a estadísticas de primer orden en el módulo de atención propuesto.

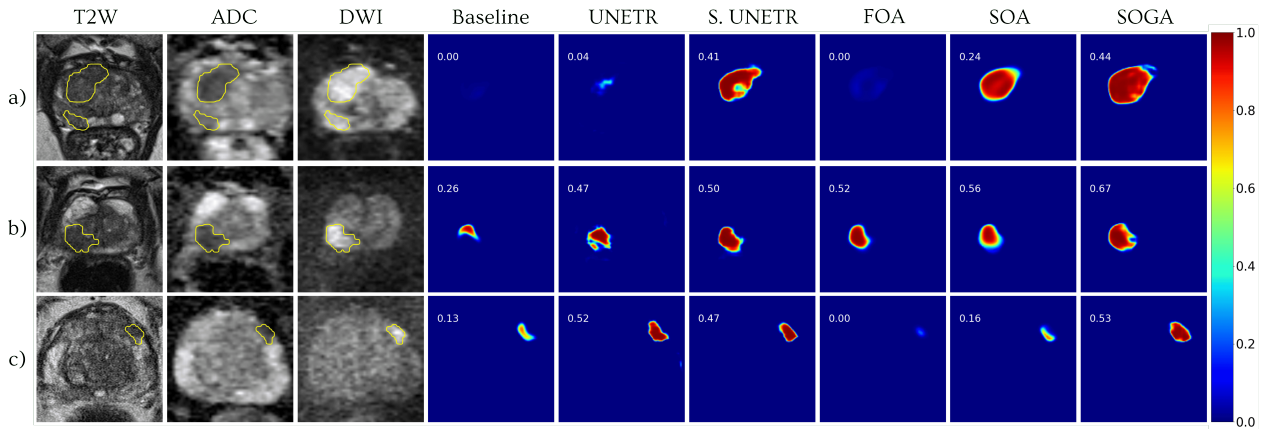


Figura 5. Visualización de predicciones para tres casos del conjunto de datos Prostate158, presentados cada uno en una fila. Las tres primeras columnas muestran imágenes bp-MRI con anotaciones de referencia (*ground-truth*). Las predicciones incluyen los valores de DSC.

En la Tabla 6 se presenta una comparación del desempeño de los distintos modelos según la zona en la que se localiza la lesión para el conjunto de datos Prostate158. Nuevamente, los modelos con atención superan de manera consistente al *baseline*, independientemente de la zona, destacándose particularmente los enfoques basados en *transformers* (UNETR y Swin UNETR). De forma análoga, los métodos fundamentados en descriptores de primer y segundo orden también muestran mejoras significativas, siendo el método propuesto, *SOGA*, el que alcanzó el mejor rendimiento en todas las métricas evaluadas. En la **zona central (CZ)**, *SOGA* incrementó el valor de AP de 0.17 a 0.59 (+247%), la sensibilidad a un falso positivo de 0.28 a 0.67 (+139%) y el DSC de 0.13 a 0.38 (+192%). En la **zona periférica (PZ)**, el modelo duplicó el AP, pasando de 0.15 a 0.30 (+100%), mientras que la sensibilidad a un falso positivo pasó de 0.21 a 0.46 (+119%) y el DSC de 0.10 a 0.26 (+160%). En la práctica clínica, esta marcada ventaja en la zona central adquiere especial relevancia, pues es precisamente en esta región donde los radiólogos enfrentan mayores

dificultades para diferenciar tejidos clínicamente significativos de cáncer de próstata ⁵

Tabla 6. Resultados en zona central (CZ) y zona periférica (PZ). Comparación entre la nnU-Net (*baseline*) y los métodos basados en atención incorporados en este framework. Las mejores métricas están resaltadas en **negrilla**.

Modelo	CZ			PZ		
	AP	Sen@1FP	DSC	AP	Sen@1FP	DSC
Baseline	0.17	0.28	0.13	0.15	0.21	0.10
UNETR	0.41	0.44	0.24	0.26	0.36	0.20
Swin UNETR	0.46	0.51	0.29	0.21	0.32	0.18
FOA	0.37	0.38	0.21	0.13	0.28	0.13
SOA	0.41	0.46	0.24	0.27	0.41	0.22
SOGA*	0.59	0.67	0.38	0.30	0.46	0.26

6. CONCLUSIONES Y TRABAJO FUTURO

Este trabajo propuso un módulo de atención geométrica utilizando descriptores de segundo orden (*SOGA*) para capturar y realzar información relevante en las conexiones de salto *skip connections* entre el codificador y el decodificador, mejorando así la segmentación y detección de lesiones de cáncer de próstata (PCa) en imágenes de *bp-MRI*. La atención geométrica propuesta aprovecha las relaciones por pares entre características a partir de descriptores de tipo *Symmetric Positive Definite* (SPD) dentro de una variedad Riemanniana y las evaluaciones demostraron la ventaja de incluir la atención propuesta en arquitecturas U-Net y nnU-Net estándar, incluso en comparación con modelos alternativos basados en atención que emplean arquitecturas *transformer*. De hecho, en el mejor de los casos, el método propuesto logró una ganancia significativa frente a las arquitecturas *transformer*, alcanzando un AUC-ROC de 0.75 frente a 0.73 (+2.7%), un AP de 0.37 frente a 0.29 (+27.6%), un DSC de 0.31 frente a 0.23 (+34.8%) y una SEN@1FP de 0.52 frente a 0.38 (+36.8%). Esto indica que los descriptores de segundo orden, junto con el procesamiento geométrico utilizado en el módulo *SOGA*, capturan de manera efectiva patrones relevantes del PCa y mejoran la robustez de las arquitecturas base para detección y segmentación.

La integración de mecanismos de atención de segundo orden sobre la nnU-Net llevó a mejoras en la capacidad de generalización con respecto a las arquitecturas base, incrementando aún más el rendimiento en datos externos, con un AUC-ROC de 0.83, Sen@1FP de 0.74 y un DSC de 0.42. De forma destacada, la ventaja del método propuesto sobre la línea base y los enfoques basados en atención fue aún más pronunciada en una cohorte completamente independiente, superando al modelo nnU-Net base en un 21.0% en AUC-ROC, 126.1% en Sen@1FP y 158.3% en DSC. Esto demuestra su capacidad para capturar características relevantes y robustas del cáncer de próstata y mantener un rendimiento sólido en nuevos escenarios, lo que resulta significativo para una futura

adopción clínica en muestras prospectivas y en diferentes centros médicos. Además, el enfoque propuesto ofrece una contribución sustancial a la detección y segmentación de lesiones pequeñas, clave en la detección temprana, uno de los mayores retos tanto para radiólogos expertos como para modelos de aprendizaje profundo. Finalmente, es de resaltar la robustez del modelo propuesto (*SOGA*), tanto para la detección como para la segmentación de lesiones pequeñas y en zonas de difícil discernimiento, como la zona central de la próstata, logrando mejoras de hasta un 61% en sensibilidad y un 52% en precisión promedio frente al segundo mejor método (*SOA*). Esto evidencia su capacidad de generalización y su relevancia clínica para identificar lesiones que podrían pasar desapercibidas con otros enfoques.

Durante el desarrollo del método propuesto *SOGA*, se observó como una limitación de estos enfoques geométricos su notable costo computacional durante el entrenamiento, asociado a operaciones sobre variedades Riemannianas, como la descomposición espectral y el mapeo bilineal que considera la geometría. Como trabajo futuro, consideramos relevante explorar la integración de *SOGA* en nuevos marcos de detección, así como en modelos fundacionales. También es interesante investigar cómo usar el análisis geométrico en variedades diferenciables alternativas a las utilizadas en el trabajo presente. En este mismo sentido, el cálculo de representaciones compactas SPD puede permitir adaptar de manera más eficiente modelos fundacionales entrenados en grandes cantidades de datos a entornos controlados, como es nuestro caso, para la segmentación y detección de lesiones csPCa. En línea con los resultados de generalización, también es importante evaluar la capacidad del módulo propuesto en escenarios con datos limitados y realizar estudios adicionales que amplíen el número de conjuntos de datos, así como continuar agregando cohortes externas e independientes. Por último, se pretende explorar en trabajos futuros diferentes alternativas de atención que permitan enfocarse no solo a nivel de canales, sino también resaltar patrones espaciales locales.

BIBLIOGRAFÍA

ADAMS, Lisa C et al. “Prostate158-An expert-annotated 3T MRI dataset and algorithm for prostate cancer detection”. In: *Computers in Biology and Medicine* 148 (2022), p. 105817 (cit. on p. 36).

ALI, Amin et al. “Prostate zones and cancer: lost in transition?” In: *Nature Reviews Urology* 19.2 (2022), pp. 101–115 (cit. on p. 19).

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. 15).

BOOTHBY, William M. *An introduction to differentiable manifolds and Riemannian geometry, Revised*. Vol. 120. Gulf Professional Publishing, 2003 (cit. on p. 31).

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 pp. 10, 26).

CAGLIC, I and BARRETT, T. *Diffusion-weighted imaging (DWI) in lymph node staging for prostate cancer. Transl Androl Urol*. 2018; 7 (5): 814–23 (cit. on p. 17).

CHAHAL, Ekam Singh et al. “Unet based xception model for prostate cancer segmentation from MRI images”. In: *Multimedia Tools and Applications* 81.26 (2022), pp. 37333–37349 (cit. on p. 23).

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. 18).

DEBS, Noëlie et al. “Evaluation of a deep learning prostate cancer detection system on biparametric MRI against radiological reading”. In: *European Radiology* (2024), pp. 1–10 (cit. on p. 24).

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. 17, 18).

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. 17).

DURAN, Audrey et al. “ProstAttention-Net: A deep attention model for prostate cancer segmentation by aggressiveness in MRI scans”. In: *Medical Image Analysis* 77 (2022), p. 102347 (cit. on pp. 11, 24).

GATTI, Marco et al. “Prostate cancer detection with biparametric magnetic resonance imaging (bpMRI) by readers with different experience: performance and comparison with multiparametric (mpMRI)”. In: *Abdominal Radiology* 44 (2019), pp. 1883–1893 (cit. on pp. 10, 11).

GROARKE, AnnMarie et al. “Quality of life and adjustment in men with prostate cancer: Interplay of stress, threat and resilience”. In: *Plos one* 15.9 (2020), e0239469 (cit. on p. 13).

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 p. 31).

HAFIZ, Abdul Mueed; PARAH, Shabir Ahmad, and BHAT, Rouf UI Alam. “Attention mechanisms and deep learning for machine vision: A survey of the state of the art”. In: *arXiv preprint arXiv:2106.07550* (2021) (cit. on p. 11).

HATAMIZADEH, Ali et al. “Swin unetr: Swin transformers for semantic segmentation of brain tumors in mri images”. In: *International MICCAI brainlesion workshop*. Springer. 2021, pp. 272–284 (cit. on p. 37).

HATAMIZADEH, Ali et al. “Unetr: Transformers for 3d medical image segmentation”. In: *Proceedings of the IEEE/CVF winter conference on applications of computer vision*. 2022, pp. 574–584 (cit. on p. 37).

HU, Jie; SHEN, Li, and SUN, Gang. “Squeeze-and-excitation networks”. In: *Proceedings of the IEEE conference on computer vision and pattern recognition*. 2018, pp. 7132–7141 (cit. on pp. 31, 34).

HUANG, Zhiwu and VAN GOOL, Luc. “A riemannian network for spd matrix learning”. In: *Proceedings of the AAAI conference on artificial intelligence*. Vol. 31. 1. 2017 (cit. on pp. 31, 32, 37).

ISENSEE, Fabian et al. “nnU-Net: a self-configuring method for deep learning-based biomedical image segmentation”. In: *Nature methods* 18.2 (2021), pp. 203–211 (cit. on pp. 21, 22, 37, 38).

JACOBS, Michael A et al. “Diffusion weighted imaging with adc mapping and spectroscopy in prostate cancer”. In: *Medicine, Baltimore* 1500 (), p. 21205 (cit. on p. 18).

KARAGOZ, Ahmet et al. “Anatomically guided self-adapting deep neural network for clinically significant prostate cancer detection on bi-parametric MRI: a multi-center study”. In: *Insights into Imaging* 14.1 (2023), p. 110 (cit. on pp. 24, 39).

KERVADEC, Hoel et al. “Beyond pixel-wise supervision for segmentation: A few global shape descriptors might be surprisingly good!” In: *Medical Imaging with Deep Learning*. PMLR. 2021, pp. 354–368 (cit. on p. 30).

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. 17).

KUNKEL, Elisabeth JS et al. “Biopsychosocial aspects of prostate cancer”. In: *Psychosomatics* 41.2 (2000), pp. 85–94 (cit. on p. 13).

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. 18).

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. 16).

LI, Wenhao et al. “Adaptive window adjustment with boundary DoU loss for cascade segmentation of anatomy and lesions in prostate cancer using bpMRI”. In: *Neural Networks* 181 (2025), p. 106831 (cit. on p. 24).

LI, Yuheng et al. “Cross-shaped windows transformer with self-supervised pretraining for clinically significant prostate cancer detection in bi-parametric MRI”. In: *Medical Physics* 52.2 (2025), pp. 993–1004 (cit. on p. 24).

LOEB, Stacy et al. “Systematic review of complications of prostate biopsy”. In: *European urology* 64.6 (2013), pp. 876–892 (cit. on p. 14).

MAURER, Martin H and HEVERHAGEN, Johannes T. “Diffusion weighted imaging of the prostate—principles, application, and advances”. In: *Translational andrology and urology* 6.3 (2017), p. 490 (cit. on pp. 17, 18).

MURPHY, Gillian et al. “The expanding role of MRI in prostate cancer”. In: *American Journal of Roentgenology* 201.6 (2013), pp. 1229–1238 (cit. on p. 16).

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. 10, 14).

OKTAY, Ozan et al. “Attention u-net: Learning where to look for the pancreas”. In: *arXiv preprint arXiv:1804.03999* (2018) (cit. on pp. 29, 30).

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. 13).

REBELLO, Richard J et al. “Prostate cancer”. In: *Nature Reviews Disease Primers* 7.1 (2021), p. 9. DOI: [10.1038/s41572-020-00243-0](https://doi.org/10.1038/s41572-020-00243-0) (cit. on p. 14).

REY, Rodolfo Alberto et al. “Male hypogonadism: an extended classification based on a developmental, endocrine physiology-based approach”. In: *Andrology* 1.1 (2013), pp. 3–16 (cit. on p. 13).

RONNEBERGER, Olaf; FISCHER, Philipp, and BROX, Thomas. “U-net: Convolutional networks for biomedical image segmentation”. In: *Medical image computing and computer-assisted intervention—MICCAI 2015: 18th international conference, Munich, Germany, October 5-9, 2015, proceedings, part III 18*. Springer. 2015, pp. 234–241 (cit. on pp. 11, 19, 21, 23, 29).

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. 11).

SAHA, Anindo et al. “Artificial intelligence and radiologists in prostate cancer detection on MRI (PI-CAI): an international, paired, non-inferiority, confirmatory study”. In: *The Lancet Oncology* 25.7 (2024), pp. 879–887 (cit. on pp. 35, 36, 38).

SAHA, Anindo et al. “The PI-CAI challenge: public training and development dataset”. In: *Zenodo, Jun* (2022) (cit. on p. 40).

SÁNCHEZ-ORO, R et al. “Prostate MRI: practical guidelines for interpreting and reporting according to PI-RADS version 2.1”. In: *Radiología (English Edition)* 62.6 (2020), pp. 437–451 (cit. on p. 15).

SÁNCHEZ-ORO, Raquel 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. 16).

SCIALPI, Michele et al. “Biparametric MRI of the prostate”. In: *Turkish journal of urology* 43.4 (2017), p. 401 (cit. on pp. 14, 15).

SHOAG, Jonathan and BARBIERI, Christopher E. “Clinical variability and molecular heterogeneity in prostate cancer”. In: *Asian journal of andrology* 18.4 (2016), pp. 543–548 (cit. on pp. 11, 48).

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. 14).

SMITH JR, Joseph A; SOLOWAY, Mark S, and YOUNG, Marvin J. “Complications of advanced prostate cancer”. In: *Urology* 54.6 (1999), pp. 8–14 (cit. on p. 13).

SYER, Tom J et al. “The diagnostic accuracy of high b-value diffusion-and T 2-weighted imaging for the detection of prostate cancer: a meta-analysis”. In: *Abdominal Radiology* 43 (2018), pp. 1787–1797 (cit. on p. 17).

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. 14).

THOMPSON, James et al. “The role of magnetic resonance imaging in the diagnosis and management of prostate cancer”. In: *BJU international* 112 (2013), pp. 6–20 (cit. on pp. 10, 11, 16).

VASWANI, Ashish et al. “Attention is all you need”. In: *Advances in neural information processing systems* 30 (2017) (cit. on p. 30).

WALMSLEY, Lee Anne. *The psychological well-being of men diagnosed with prostate cancer*. University of Kentucky, 2015 (cit. on p. 13).

WEI, Chao et al. “Enhancing prostate cancer segmentation in bpMRI: Integrating zonal awareness into attention-guided U-Net”. In: *Digital Health* 11 (2025), p. 20552076251314546 (cit. on pp. 23, 25).

WOO, Sanghyun et al. “Cbam: Convolutional block attention module”. In: *Proceedings of the European conference on computer vision (ECCV)*. 2018, pp. 3–19 (cit. on p. 29).

XU, Helen et al. “Prostate cancer detection using residual networks”. In: *International journal of computer assisted radiology and surgery* 14 (2019), pp. 1647–1650 (cit. on p. 15).

YANG, Chunguang et al. “Deep learning model for the detection of prostate cancer and classification of clinically significant disease using multiparametric MRI in comparison to PI-RADs score”. In: *Urologic Oncology: Seminars and Original Investigations*. Vol. 42. 5. Elsevier. 2024, 158–e17 (cit. on p. 23).

LU-YAO, Grace L et al. “Outcomes of localized prostate cancer following conservative management”. In: *Jama* 302.11 (2009), pp. 1202–1209 (cit. on p. 18).

ANEXOS

Anexo A. Productos Académicos

- **Ortiz, M.**, Olmos, J., Martínez, F. (2025). A Second-Order Attention Mechanism For Prostate Cancer Segmentation and Detection in Bi-Parametric MRI. In *28th Iberoamerican Congress on Pattern Recognition (CIARP) Springer Proceedings*. Aceptado.
- **Ortiz, M.**, Olmos, J., Martínez, F. (2025). Mecanismos de atención de segundo orden para la segmentación de lesiones de cáncer de próstata en secuencias de resonancia magnética bi-paramétrica. En *Inngeniarte-c 2025, Universidad Autónoma de Bucaramanga (UNAB)*.