

ACUERDO n.º 292 DE 2023
05 de Septiembre

1

Por el cual se aprueba el otorgamiento de la distinción “*Trabajo de Grado Meritorio*” al trabajo presentado por YESID ALFONSO GUTIÉRREZ GUATE, estudiante del programa Maestría en Ingeniería de Sistemas e Informática

EL CONSEJO ACADÉMICO DE LA UNIVERSIDAD INDUSTRIAL DE SANTANDER,
en uso de sus atribuciones legales, y

CONSIDERANDO,

- a. Que según lo establecido en la Ley 30 de 1992 y el Estatuto General de la Universidad industrial de Santander, aprobado mediante acuerdo del Consejo Superior mediante Acuerdo n.º 166 del 22 de diciembre de 1993, el Consejo Académico es la máxima autoridad académica de la Universidad.
- b. Que según lo establecido en el literal b. del artículo 22, del Estatuto General, aprobado por el Consejo Superior (Acuerdo n.º 166 de 1993), el Vicerrector Académico preside el Consejo Académico en ausencia del Rector.
- c. Que el director de la Escuela de Ingeniería de Sistemas e Informática, solicitó al Consejo Académico otorgar la distinción ‘Trabajo de Grado Meritorio’ al trabajo de grado titulado “*Prostate lesions characterization in MRI sequences using a deep contrastive learning framework*”, elaborado por el estudiante Yesid Alfonso Gutiérrez Guate del programa de Maestría en Ingeniería de Sistemas e Informática y dirigido por el profesor Fabio Martínez Carrillo y codirigido de profesor John Edison Arévalo Ovalle.
- d. Que los evaluadores del trabajo de grado, profesores, Oscar Acosta Tamayo y Claudio Delrieux, recomendaron otorgar la distinción “Trabajo de Grado Meritorio” al trabajo referido en el literal c), en consideración al cumplimiento de las disposiciones contenidas en el Artículo 110 del Reglamento General de Posgrados, teniendo en cuenta que en su concepto constituye un aporte significativo en el área, concepto que hace parte integral del presente acuerdo.
- e. Que el Consejo Académico, en sesión del 05 de agosto de 2023, aprobó la solicitud referida en el considerando c), del presente acto administrativo.

En mérito de lo anterior,

ACUERDA:

ARTÍCULO 1º. Aprobar la distinción ‘Trabajo de Grado Meritorio’ al trabajo titulado “*Prostate lesions characterization in MRI sequences using a deep contrastive learning framework*”, elaborado por el estudiante Yesid Alfonso Gutiérrez Guate del programa de Maestría en Ingeniería de Sistemas e Informática y dirigido por el profesor Fabio Martínez Carrillo y codirigido de profesor John Edison Arévalo Ovalle.



ACUERDO n.º 292 DE 2023
05 de Septiembre

2

ARTÍCULO 2º. Informar sobre el contenido del presente acuerdo a la Dirección de Admisiones y Registro Académico y la Escuela de Ingeniería de Sistemas e Informática, para lo de su competencia.

PUBLÍQUESE, COMUNÍQUESE Y CÚMPLASE

Expedido en Bucaramanga, a los cinco (05) días del mes de septiembre de 2023.

EL PRESIDENTE DEL CONSEJO ACADÉMICO,


DANIEL ALFONSO SIERRA BUENO
Vicerrector Académico

LA SECRETARIA GENERAL,


SOFÍA PINZÓN DURÁN

Bucaramanga, 23 de agosto de 2023.



Señores
CONSEJO ACADÉMICO
Universidad Industrial de Santander
Bucaramanga

ASUNTO: Distinción Trabajo de Grado Meritorio para el Trabajo de Investigación
Maestría en Ingeniería de Sistemas e Informática.

Cordial saludo:

Por medio de la presente comunicación, me permito enviar la recomendación del jurado evaluador para que se otorgue la distinción **Trabajo de Grado Meritorio** al Trabajo de Investigación titulado **“Prostate lesions characterization in MRI sequences using a deep contrastive learning framework”**; realizado por el estudiante de Maestría en Ingeniería de Sistemas e Informática Yesid Alfonso Gutiérrez Guate, con código de estudiante **2208454**. Este Trabajo de investigación cumple con los requisitos establecidos en el Artículo 110 del Acuerdo del Consejo Superior No. 075 de 2013 (Reglamento General de Posgrado). Con respecto al párrafo de este artículo, el estudiante y autor del Trabajo de Investigación inició sus estudios de Maestría en el **segundo periodo académico del 2020**, culminó las asignaturas del programa y su Trabajo de Investigación en el primer semestre académico del 2023 siendo este su sexto (6) periodo académico. El estudiante entregó el 30 de marzo de 2023 el informe correspondiente a su Trabajo de Investigación para su respectiva evaluación. El cual recibió la calificación de **APROBADO** el **13 de julio de 2023** por parte del jurado calificador.

Los resultados reportados en el presente Trabajo de Investigación:

- Gutiérrez, Y., Arevalo, J., & Martínez, F. (2022, July). “Multimodal Contrastive Supervised Learning to Classify Clinical Significance MRI Regions on Prostate Cancer”. In *2022 44th Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC)* (pp. 1682-1685). IEEE.
Artículo publicado en Revista categoría B según Publindex (ISSN 2694-0604).
[Paper](#)



- Gutiérrez, Y., Arevalo, J., & Martínez, F. (2022). An inception-based deep multiparametric net to classify clinical significance MRI regions of prostate cancer. *Physics in Medicine & Biology*, 67(22), 225004.

Artículo publicado en Revista categoría A1 según Publindex (ISSN 0031-9155)

[Paper](#)

A esta comunicación se adjunta el concepto del jurado calificador.

Atentamente,



FABIO MARTÍNEZ CARRILLO
Coordinador de los Programas de Posgrados

TRABAJO DE GRADO MERITORIO		TRABAJO DE GRADO LAUREADO			
<p>a). Según el concepto del jurado, el trabajo constituya un aporte al campo disciplinar en que se desarrolle el trabajo de grado. <i>Artículo 110 literal a). Reglamento General de Posgrados Acuerdo 075 de 2013 del Consejo Superior. Por favor escribir el aporte científico.</i></p> <p>El maestrando realizó un trabajo de investigación original, que tiene en cuenta el estado del arte y genera un aporte significativo al mismo. Tanto la disertación como la presentación se presentaron de una manera clara y bien estructurada. Se consiguieron los objetivos planteados, y los resultados constituyen un avance claro al estado del arte.</p>		<p>a). Según el concepto del jurado, el trabajo constituya un aporte al campo disciplinar en que se desarrolle el trabajo de grado. Artículo 110 literal a). Reglamento General de Posgrados Acuerdo 075 de 2013 del Consejo Superior. Por favor escribir el aporte científico.</p>			
<p>b). Publicación o aceptación de (1) artículo de su autoría, en revistas indexadas u homologadas por COLCIENCIAS en categoría A o B según la clasificación vigente de PUBLINDEX, de COLCIENCIAS, o en revistas con índice de impacto equivalentes a estas categorías, que contengan expresamente los avances o resultados del trabajo de grado. Artículo 110 literal c). Reglamento General de Posgrados Acuerdo 075 de 2013 del Consejo Superior.</p>		<p>SI <input checked="" type="checkbox"/> NO <input type="checkbox"/></p>	<p>b). Publicación o aceptación de (2) artículo de su autoría, en revistas indexadas u homologadas por COLCIENCIAS en categoría A según la clasificación vigente de PUBLINDEX, de COLCIENCIAS, o en revistas con índice de impacto equivalentes a estas categorías, que contengan expresamente los avances o resultados del trabajo de grado. Artículo 110 literal b). Reglamento General de Posgrados Acuerdo 075 de 2013 del Consejo Superior.</p>	<p>SI <input type="checkbox"/> NO <input type="checkbox"/></p>	
<p>c). Participación con ponencia en, al menos, un (1) evento académico internacional. Esta ponencia debe incluir expresamente los avances o resultados del trabajo de grado. <i>Artículo 110 literal c). Reglamento General de Posgrados Acuerdo 075 de 2013 del Consejo Superior.</i></p>		<p>SI <input checked="" type="checkbox"/> NO <input type="checkbox"/></p>	<p>c). Participación con ponencia en, al menos, un (1) evento académico internacional. Esta ponencia debe incluir expresamente los avances o resultados del trabajo de grado. Artículo 110 literal c). Reglamento General de Posgrados Acuerdo 075 de 2013 del Consejo Superior.</p>		<p>SI <input type="checkbox"/> NO <input type="checkbox"/></p>
<p>d). Para acceder a las distinciones de Trabajo de grado laureado y Trabajo meritorio, el estudiante-autor del trabajo de grado debe haber finalizado el programa, en su totalidad, en el tiempo establecido por la universidad para ello. <i>Artículo 139 parágrafo 1. Reglamento General de Posgrados Acuerdo 075 de 2013 del Consejo Superior.</i></p>		<p>SI <input checked="" type="checkbox"/> NO <input type="checkbox"/></p>	<p>d). Para acceder a las distinciones de Trabajo de grado laureado y Trabajo meritorio, el estudiante-autor del trabajo de grado debe haber finalizado el programa, en su totalidad, en el tiempo establecido por la universidad para ello. <i>Artículo 139 parágrafo 1. Reglamento General de Posgrados Acuerdo 075 de 2013 del Consejo Superior.</i></p>		<p>SI <input type="checkbox"/> NO <input type="checkbox"/></p>
<p>Consideraciones, Comentarios y Observaciones:</p> <p>El trabajo de investigación propuesto supera los alcances académicos en un trabajo de maestría, fundamentados en los productos académicos alcanzados. Además, los objetivos fueron plenamente cumplidos, así como el modelo de segmentación propuesto tiene resultados competitivos en el estado del arte</p>					

NOMBRE: Claudio Delrieux

FIRMA:



Fecha: 15 de julio de 2023



44th Annual International Conference of
**the IEEE Engineering in
Medicine and Biology Society**

11-15 July 2022  Glasgow, Scotland, UK



Certificate of Participation

awarded to

Yesid Gutierrez

*For the participation in the
44th International Engineering in Medicine and Biology
Conference (EMBC'22)
11-15 July 2022, Glasgow, Scotland, UK*

The Institute of Electrical and Electronic Engineers
Engineering in Medicine and Biology Society

www.embs.org | www.IEEE.org

Multimodal Contrastive Supervised Learning to Classify Clinical Significance MRI Regions on Prostate Cancer

Yesid Gutiérrez¹ John Arevalo², and Fabio Martínez³

Abstract—Clinically significant regions (CSR), captured over multi-parametric MRI (mp-MRI) images, have emerged as a potential screening test for early prostate cancer detection and characterization. These sequences are able to quantify morphology, micro-circulation, and cellular density patterns that might be related to cancer disease. Nonetheless, this evaluation is mainly carried out by expert radiologists, introducing inter-reader variability in the diagnosis. Therefore, different deep learning models were proposed to support the diagnosis, but a proper representation of prostate lesions remains limited due to the non-alignment among sequences and the dependency of considerable amounts of labeled data for learning. The main limitation of such representation lies in the cross-entropy minimization that only exploits inter-class variation, being insufficient data augmentation and transfer learning strategies. This work introduces a Supervised Contrastive Learning (SCL) strategy that fully exploits the inter and intra-class variability of prostate lesions to robustly represent MRI regions. This strategy extracts lesion sample tuples, with positive and negative labels, regarding a query lesion. Such tuples are involved into an easy-positive, and semi-hard negative mining to project samples that better update the deep representation. The proposed learning strategy achieved an average ROC-AUC of 0.82, to characterize prostate cancer in MRI, using only the 60% of the available annotated data.

Clinical relevance— A robust learning scheme that properly finds representations in limited data scenarios to classify clinically significant MRI regions on prostate cancer.

I. INTRODUCTION

The American Cancer Society reported 191.930 new cases of prostate cancer in the United States during 2020, becoming the cancer with highest incidence in men [1]. The early detection and diagnosis of this disease may significantly reduce deaths. However, current over-diagnosis limits a proper characterization of malignant tumors and treatment of the disease [1]. Nowadays, the Prostatic Specific Antigen (PSA) is the most common screening test at early stages but with significant low specificity (approx. 25%), requiring additional tests to guarantee an effective and correct diagnosis [2]. Also, the digital rectal examination is highly dependent of physician’s experience and only effective to detect malignant tumors at peripheral prostate zone. Currently, a promising alternative to support the early diagnosis and screening of prostate cancer lies on the analysis of multi-parametric

Magnetic Resonance Imaging (mp-MRI) sequences. Among others, these sequences estimate in-vivo the aggressiveness of malignant tumors, even located far from the rectum wall [3], [4]. Among the most notable sequences are the T2-weighted imaging (T2WI), the Diffusion Weighted Imaging (DWI), and the Dynamic Contrast Enhanced (DCE). Particularly, T2WI is used to detect and grade cancer, as well as to characterize the morphological features of the prostatic gland. Complementary, DWI allows to quantify the cellular density of the tissues, being the ADC, and the Maximum B value (B-VAL) images, remarkable modalities from such sequences. Also, K^{trans} maps from DCE sequences support quantification of the tumor aggressiveness, as well as the recognition of micro-circulation properties, reflecting vascular patterns as a response of gadolinium influx [5], [6]. Nevertheless, the analysis of prostate cancer lesions is commonly based on radiologist experience, introducing inter-reader variability and over-diagnosis, resulting in a uncontrolled progression of the disease.

Nowadays, deep learning strategies have been proposed to quantitatively support lesion analysis, to enhance cancer characterization and to achieve a relationship of radiological findings with biopsy results. For instance, Mehrtaash *et. al* proposed a multimodal 3D CNN architecture that integrated regions of ADC, B-VAL, and K^{trans} images with zonal information of the prostatic gland [7]. Also, Chen *et. al* proposed a multimodal VGG-16 architecture that integrated T2WI, ADC and K^{trans} maps into an RGB channel scheme using a transfer learning (TL) strategy from ImageNet dataset [8]. Afterwards, Bleker *et. al* extracted 92 radiomic features from the available mp-MRI data [9]. These radiomic features were used as descriptor of the lesions using a Random forest (RF) and an extreme gradient boosting (XGB) algorithm. Also, Aldojo *et. al*, proposed a 3D mp-MRI multi-channel CNN architecture that uses different combinations of T2WI, DWI, ADC and K^{trans} images [10]. These multimodal deep strategies have evidenced a potential support to represent, characterize and reduce false positives of malignant prostate lesions. Nonetheless, these representations remain dependent of a stratified, balanced, and well-labeled mp-MRI data, which may be a critical issue on clinical scenarios. Despite that some studies have addressed this issue from data augmentation [7], [9], [10] and transfer learning techniques [8], the representations remain insufficient to generalize explicit learned information. This fact may be associated to a cross-entropy rule that only follows an inter-class minimization without considering high self class variability.

¹ Yesid Gutiérrez is a member of Biomedical Imaging Vision and Learning Laboratory (BIVL²ab), Universidad Industrial de Santander, Colombia. yesid.gutierrez@saber.uis.edu.co,

²John Arevalo is a member of Machine Learning Analysis and computer vision (M_L ACV), Universidad Industrial de Santander, Colombia. jearevalo@unal.edu.co

³Fabio Martínez is with the Department of Systems Engineering, Universidad Industrial de Santander, Colombia. famarc@uis.edu.co

The main contribution of this work is a supervised contrastive learning (SCL) strategy that fully exploits positive and negative visual samples, allowing to discriminate CSR over MRI-MP sequences. The proposed strategy extends traditional learning algorithms, by including an energy based learning scheme that models the inter and intra-variability of prostate lesions represented as textural similarities among CSR. We selected a state-of-the-art backbone that not only characterizes malignant prostate lesions from mp-MRI sequences, but also discriminates such lesions according to the available zonal information [7]. Then an easy-positive mining strategy was herein implemented to find tuples of training sample configurations that better update the deep representation. Then, a contrastive loss was implemented to build an embedding space where distance among query and negative lesions were maximized. Finally, The proposed learning strategy was validated in three sub-sampling schemes to emulate different challenging clinical scenarios. The achieved results show that the proposed SCL strategy obtained a better performance even in smaller sub-data sets. Hence, the proposed representation is able to learn inter and intra-class variability, exploiting textural similarities among the annotated data.

II. PROPOSED APPROACH

This study proposes a SCL strategy to fit a multimodal deep representation to classify CSR in MRI under limited availability of annotated data scenarios. The general pipeline of the proposed scheme is illustrated in Figure 1.

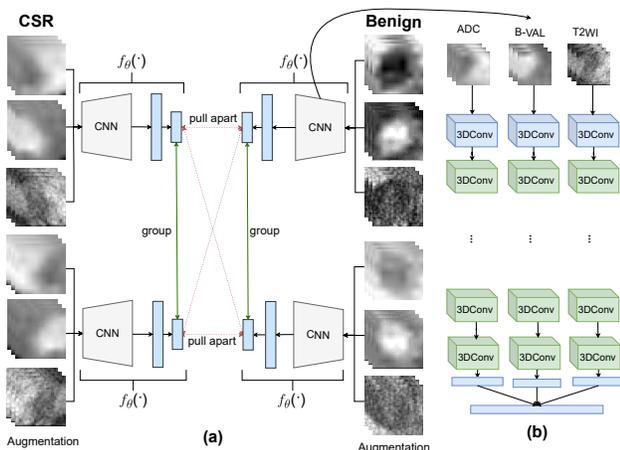


Fig. 1. Pipeline of the proposed approach: (a) each prostate lesion is augmented and projected into an embedding vector. Then, these embedding vectors are grouped or pulled off depending on their categories by using a contrastive loss function. (b) A state-of-the-art visual representation that integrates the available mp-MRI data [7].

A. Visual representation of prostate lesions in MRI

This work uses as deep backbone representation, a multimodal convolutional architecture inspired in Mehrtaash *et al.* [7], in the same context of classification of CSR in MP-MRI images. The original network uses as inputs the ADC, B-VAL and K^{trans} images. Nonetheless, current studies in

clinical practice suggest that DCE-MRI may not be relevant to characterize malignant lesions, while the T2WI images allow to localize and stage tumors [5], [11]. Following these assumptions and the potential usability in clinical routine, we decided to perform a bp-MRI approach that integrates ADC, B-VAL, and T2WI maps to characterize CSR. Then as seen in Figure 1-b, the multimodal network takes each MRI modality and models 3D convolutional branches independently. The resultant last-embedding vectors of each modality are fused following a concatenation scheme. Additionally, the zonal information is integrated in the fused representation to include information about disease occurrence. It is noteworthy that the proposed SCL approach has the flexibility of any deep representation that characterizes CSR in MRI.

B. Learning lesions from contrastive representations

For supervised contrastive learning, the backbone f_θ (a convolutional model with θ parameters) is updated during each epoch by maximizing an embedding space that represents positive and negative prostate regions. Firstly, the selected backbone is modified to project a set of embedding vectors that represents the multimodal input of prostate regions. Then a learning scheme is defined to build an optimal embedding space that maximizes the distance among different classes (inter class-variability), while clustering lesions of same class (intra-class variability). Each MP-MRI regions x_k together with zonal information, are projected to the convolutional representation f_θ to obtain the respective embedding vector $f_\theta(x_k) \in \mathbb{R}^n$. These Euclidean vectors are then normalized to form a unit hyper-sphere allowing a computational efficiency during training.

During the training phase, each batch of data is composed by query-positive lesion pairs (x_i, x_j) and a set of negative lesion samples $\{x_k\}$. This sample is approached from the same image data by following image augmentations on each of the lesions. These new samples, obtained from transformations, introduce variability on the data distribution allowing to properly measure prostate lesion similarities between query-positive lesions and textural variability among query-negative lesions. Here, the embedding space is built from a contrastive learning rule, implemented as the NT-Xent contrastive loss objective function [12], defined as:

$$L_{i,j} = -\log \frac{\exp(f_\theta(x_i)^\top \cdot f_\theta(x_j)/\tau)}{\sum_{k=1}^{2N} \mathbb{1}_{[k \neq i]} \exp(f_\theta(x_i)^\top \cdot f_\theta(x_k)/\tau)}$$

Where τ is a temperature hyper-parameter that controls the sensitivity of the similarity measurements. This objective loss function takes advantage of positive and negative lesions to better exploit inter and intra class variability, being robust to obtain good representations from limited number of annotated MRI data. The end-to-end classification may lose topological properties from the embedded representation, such as, the unit norm, the distance, and the angle between vectors to describe malignant lesions. Therefore, the resulting contrastive space is used as a descriptor of prostate cancer. Finally, a Logistic Regression classifier was selected

to linearly discriminate the embedded vectors as malignant (CSR) or benign according to the hyper-plane separation.

C. Contrastive lesion mining

Contrastive learning approaches require to properly select tuples that are sufficiently challenging to variate the representation to capture main textural moments that discriminate among lesions. Particularly, in any dataset with a total of N samples, the contrastive learning may take $2N$ annotated lesions, considering the augmentation of samples from simple transformations. In such case, the total amount of possible tuples is $8N^3$, being many of these tuples noise and sub-optimal to help regarding the task of updating the convolutional representation.

This work considered the query-positive (x_i, x_j) and query-negative (x_i, x_k) tuples to map relevant tuples and guarantee a convergence of the learning strategy. The easy-positive mining strategy finds the most similar positive sample x_j in the mp-MRI image x with respect to the same label $C(x_i) = C(x)$ as the anchor sample x_i . This search is formally defined as $x_j = \arg \min_{x: C(x)=C(x_i)} d(f(x_i), f(x))$, where $d(\cdot)$ is the euclidean distance among two given lesions. Complementary, the semi-hard negative mining finds the most significant negative lesion x_k , where $x_k = \arg \min_{x: C(x) \neq C(x_i)} d(f(x_i), f(x))$, constrained by $d(f(x_i), f(x)) > d(f(x_i), f(x_j))$. This mining strategy filters tuples for mapping during training, maintaining the intra-class variance, and producing useful gradients during the learning stage [13].

D. Experimental setup

Dataset. This public dataset was used in the PROSTATEx grand challenge, and contains 204 MRI studies which correspond to 320 labeled prostate lesions. Each case consists of four sets of MRI scan data: two sets of T2-weighted images, ADC images computed from DWI, and K^{trans} images (computed from dynamic contrast-enhanced (DCE) images) [14]. Additionally, the dataset includes annotations related with the lesion location $[k, j, i]$, the clinical significance of the lesion supported by biopsy, and the prostate zone of the study. The finding information was provided by two expert radiologists [14].

Method setup. We extracted volumetric MRI ROIs x_i centered on the radiologist annotations with a fixed voxel region size of $12 \times 32 \times 32$ to locally represent the lesion. Each sample was artificially augmented following transformations, such as: vertical and horizontal translations, random rotations and flipping w.r.t the horizontal plane. At the training phase, we emulated three challenging scenarios with scarce of labeled lesions (benignant, stratified, and malignant lesions sub-sampling) to study the diagnosis performance of our proposed approach. Additionally, each prostate lesion ROI x_i was projected into an embedding feature vector $f_\theta(x_i) \in \mathbf{R}^{128}$, where f_θ is an encoding network. Moreover, such deep convolutional representation was optimized with a RMSprop algorithm using a momentum of 0.6 and a learning rate of 1×10^{-6} . For validation of the proposed approach, each of

the configurations was mapped to the test set defined by the authors of the challenge, returning the corresponding Area Under the Receiver Operating Characteristic Curve (ROC-AUC) to measure the performance of the diagnosis [14].

III. EVALUATION AND RESULTS

The availability of large amounts of MP-MRI annotated data is a principal limitation on clinical scenarios to implement deep learning strategies. Then, a main issue is to obtain reliable representations to support classification of CSR with relative few information. In this work, we emulate three challenging scenarios with incremental data to analyze the behavior of our SCL strategy. This learning scheme uses a multimodal network as backbone, that receives as input a bp-MRI approach that integrates prostate zonal information, ADC, B-VAL, and T2WI maps. Additionally, we selected the binary cross-entropy (BCE) as baseline to compare achieved results by the proposed training strategy.

TABLE I
TEST ROC-AUC PERFORMANCE ACHIEVED BY DIFFERENT
CONTRASTIVE CONFIGURATIONS IN A STRATIFIED SUB-SAMPLING
SCHEME

Percentage	NT-Xent-LM	NT-Xent	Triplet-LM	Triplet
60%	0.82	0.79	0.79	0.78
80%	0.81	0.81	0.74	0.75
100%	0.82	0.79	0.76	0.73

In our first experiment, we considered an ablation study to validate the contribution of different components in the proposed learning scheme, reporting the ROC-AUC performance in a stratified sub-sampling scheme. In such sense, we are interested in quantifying the contribution of NT-Xent loss w.r.t the triplet version that only considers one negative sample in the tuple. Also, we are interested in measuring the contribution of lesion mining (LM) for the proposed SCL scheme. Table I summarizes the achieved results for the different configurations. It should be noted that triplet loss is relatively independent to the mining process but with a remarkable low performance with respect to the NT-Xent loss. The best configuration was achieved with the NT-Xent loss but with a proper tuple selection, achieved by the lesion mining process.

Then, in a second experiment we compared the best contrastive configuration w.r.t the BCE learning scheme under different sub-sampling configurations. Figure 2 summarizes the achieved results for both learning schemes in the test set, showing a convincing advantage of SCL at different sub-sets of annotated lesions. These results evidence that from 60% of the annotated lesions, it is enough to obtain appropriated representations of CSR. Remarkably, our SCL scheme achieved a ROC-AUC of 0.82 in the test set using only the 60% of the annotated data, while for the baseline BCE learning, the deep representation obtained a ROC-AUC of 0.77. Moreover, Mehrtash *et al.* [7] reported a ROC-AUC of 0.80, but using K^{trans} , ADC, and B-VAL maps, and all of the available annotated lesions. This fact clearly

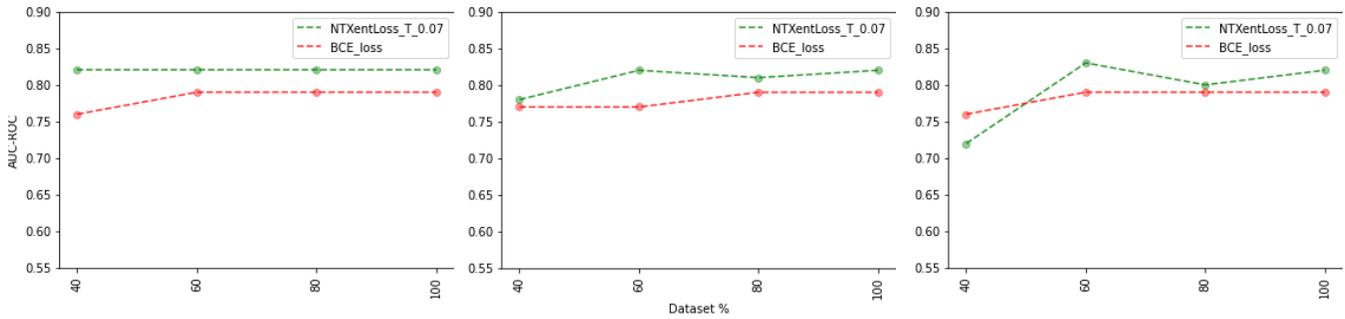


Fig. 2. ROC-AUC performance archived by Mehrtash’s architecture in the test set using a SCL strategy (green), and the classical BCE learning scheme (red) under different sub-sampling configurations. From left to right: benignant, stratified and malignant sub-sampling strategies respectively.

highlights the potential of SCL to update deep convolutional representations in clinical scenarios with scarce of data. The obtained results may be attributed to the contrastive self-entropy learning of multimodal lesions, which exploits intra and inter variabilities of prostate tissues. Besides, this SCL scheme is in general robust to imbalance data problems, a real issue on clinical scenarios. Finally, the Mann-Whitney U ranked test was estimated using both learning schemes, suggesting with a confidence of 95% (p-value=0.0001), that these learning schemes are statistically different in their probability distributions.

IV. CONCLUSIONS

This work presented a supervised deep contrastive learning strategy to describe morphological and cellular density patterns on MRI, which are related with prostate cancer disease. Results showed that SCL strategy improved the characterization of malignant lesions w.r.t supervised learning schemes based on the binary cross entropy loss. This learning strategy trained deep learning models with limited annotated MRI data, a common issue in clinical scenarios. Experimental evaluation shows that a bp-MRI approach using only the 60% of the data is able to obtain good representations of malignant prostate lesions. The validation also suggests the contrastive learning scheme better exploits similarities between prostate lesion tissues by using the annotated data more efficiently. Future work includes the validation of weakly contrastive learning to exploit textural visual information of prostate lesions.

ACKNOWLEDGMENT

The authors acknowledge the Vicerrectoría de Investigación y Extensión (VIE) of the Universidad Industrial de Santander for supporting this research work by the project: “Predicción de patologías cardiacas utilizando representaciones de aprendizaje profundo en secuencias de resonancia magnetica cardiaca (CMR)” with SIVIE code 2703.

COMPLIANCE WITH ETHICAL STANDARDS

This research study was conducted retrospectively using human subject data made available in open access by the

2017 SPIE Medical Imaging Symposium on the SPIE-AAPM-NCI Prostate MR Classification Challenge [14]. Ethical approval was not required as confirmed by the license attached with the open access data.

REFERENCES

- [1] R. L. Siegel *et al.*, “Cancer statistics, 2020,” *CA: A Cancer Journal for Clinicians*, vol. 70, no. 1, pp. 7–30, 2020.
- [2] M. J. Barry, “Prostate-specific-antigen testing for early diagnosis of prostate cancer,” *New England Journal of Medicine*, vol. 344, no. 18, pp. 1373–1377, 2001.
- [3] B. J. Manley *et al.*, “Prostate mri: a national survey of urologist’s attitudes and perceptions,” *International braz j urol*, vol. 42, no. 3, pp. 464–471, 2016.
- [4] M. Saar *et al.*, “Current role of multiparametric mri and mri targeted biopsies for prostate cancer diagnosis in germany: a nationwide survey,” *Urologia Internationalis*, vol. 104, no. 9–10, pp. 731–740, 2020.
- [5] G. Murphy, et. al, “The expanding role of mri in prostate cancer,” *American Journal of Roentgenology*, vol. 201, no. 6, pp. 1229–1238, 2013.
- [6] C. Cuenod and D. Balvay, “Perfusion and vascular permeability: basic concepts and measurement in dce-ct and dce-mri,” *Diagnostic and interventional imaging*, vol. 94, no. 12, pp. 1187–1204, 2013.
- [7] A. Mehrtash *et al.*, “Classification of clinical significance of mri prostate findings using 3d convolutional neural networks,” in *Medical Imaging 2017: Computer-Aided Diagnosis*, vol. 10134, p. 101342A, International Society for Optics and Photonics, 2017.
- [8] Q. Chen *et al.*, “A transfer learning approach for classification of clinical significant prostate cancers from mpMRI scans,” in *Medical Imaging 2017: Computer-Aided Diagnosis*, vol. 10134, p. 101344F, International Society for Optics and Photonics, 2017.
- [9] J. Bleker *et al.*, “Multiparametric mri and auto-fixed volume of interest-based radiomics signature for clinically significant peripheral zone prostate cancer,” *European radiology*, vol. 30, no. 3, pp. 1313–1324, 2020.
- [10] N. Aldoj *et al.*, “Semi-automatic classification of prostate cancer on multi-parametric mr imaging using a multi-channel 3d convolutional neural network,” *European radiology*, vol. 30, no. 2, pp. 1243–1253, 2020.
- [11] J. O. Barentsz *et al.*, “Esur prostate mr guidelines 2012,” *European radiology*, vol. 22, no. 4, pp. 746–757, 2012.
- [12] K. Sohn, “Improved deep metric learning with multi-class n-pair loss objective,” in *Advances in neural information processing systems*, pp. 1857–1865, 2016.
- [13] H. Xuan et. al, “Improved embeddings with easy positive triplet mining,” in *Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision*, pp. 2474–2482, 2020.
- [14] G. Litjens *et al.*, “Computer-aided detection of prostate cancer in mri,” *IEEE transactions on medical imaging*, vol. 33, no. 5, pp. 1083–1092, 2014.

An inception-based deep multiparametric net to classify clinical significance MRI regions of prostate cancer

Yesid Gutiérrez^{1,2}, John Arevalo², Fabio Martínez^{1,2}

¹ BivL²ab, Universidad Industrial de Santander, Bucaramanga, Colombia

² MACV, Universidad Industrial de Santander, Bucaramanga, Colombia

E-mail: famarc@uis.edu.co

Abstract. Objective: Multi-parametric magnetic resonance imaging (MP-MRI) has played an important role in prostate cancer diagnosis. Nevertheless, in the clinical routine, these sequences are principally analyzed from expert observations, which introduces an intrinsic variability in the diagnosis. Even worse, the isolated study of these MRI sequences trend to false positive detection due to other diseases that share similar radiological findings. Hence, the main objective in this study was to design, propose and validate a deep multimodal learning framework to support MRI-based prostate cancer diagnosis using cross-correlation modules that fuse MRI regions, coded from independent MRI parameter branches.

Approach: This work introduces a multimodal scheme that integrates MP-MRI sequences and allows to characterize prostate lesions related to cancer disease. For doing so, potential 3D regions were extracted around expert annotations over different prostate zones. Then, a convolutional representation was obtained from each evaluated sequence, allowing a rich and hierarchical deep representation. Each convolutional branch representation was integrated following a special inception-like module. This module allows a redundant non-linear integration that preserves textural spatial lesion features and could obtain higher levels of representation.

Main results: This strategy enhances micro-circulation, morphological, and cellular density features, which thereafter are integrated according to an inception late fusion strategy, leading to a better differentiation of prostate cancer lesions. The proposed strategy achieved a ROC-AUC of 0.82 over the PROSTATEx dataset by fusing regions of K^{trans} and Apparent Diffusion Coefficient (ADC) maps coded from DWI-MRI.

Significance: In this study was conducted an evaluation about how MP-MRI parameters can be fused, through a deep learning representation, exploiting spatial correlations among multiple lesion observations. The strategy, from a multimodal representation, learns branches representations to exploit radio-logical findings from ADC and K^{trans} . Besides, the proposed strategy is very compact (151,630 trainable parameters). Hence, the methodology is very fast in training (3 seconds for an epoch of 320 samples), being potentially applicable in clinical scenarios.

Keywords: MRI, prostate cancer, K^{trans} , multimodal learning, deep representations, inception-multimodal layers

Submitted to: *Phys. Med. Biol.*

1. Introduction

According to the American Cancer Society (ACS), 191.930 new cases and 33.330 deaths were reported due to prostate cancer in the United States during 2020, hitting a new record in deaths for the last two decades [1]. The early prostate cancer diagnosis is fundamental to properly planning treatments, consequent with a mortality reduction. Currently, diagnostic methods range from blood test to quantify the concentration of Prostate Specific Antigen glycoprotein (PSA) to the Digital Rectal Examination (DRE). Nevertheless, these classical tests have demonstrated: low specificity [2] (mainly for PSA test), and inter-examiner variability on diagnosis (for DRE) [3]. More sophisticated and accurate alternatives for diagnosis include the the trans-rectal ultrasound-guided biopsy (TRUS) [4], but with a significant reported rate of false negative cases (around $\sim 30\%$), and side effects such as rectal bleeding, bacteriuria, and sepsis [5, 6].

Multi-parametric magnetic resonance imaging (MP-MRI) has become a powerful medical diagnostic tool to identify clinically significant prostate lesions, and to confirm the presence of the disease [7, 8]. In current clinical protocols, the analysis and observation of integrated MP-MRI sequences have allowed to approximate detection and to estimate *in vivo* the cancer aggressiveness, approximating the Gleason score group [9, 10]. For instance, T2-weighted MRI sequences are useful to observe morphological prostate patterns but with predominant limitations regarding sensitivity and specificity to characterize clinically significant lesions [8]. Complementary, the DCE-MRI (Dynamic Contrast Enhanced) sequences support the characterization of suspicious lesions by providing micro-circulation parameters of tissues [11]. Among these DCE sequences, the pharmaco-kinetic features coded, such as the volume transfer constant (K^{trans}), exposes micro-circulation patterns that correlate with the non-controlled formation of blood vessels in the prostate gland (prostate Angiogenesis) [12]. Despite of the contribution of MP-MRI in the characterization of clinically significant lesions, these MRI techniques are commonly analyzed independently based only on expert observations, introducing high variability on the diagnosis and leading to false positive detection with respect to other diseases such as Benign Prostatic Hyperplasia (BPH), bacteriuria or Angiogenesis [13, 14].

Hence, computer-aided diagnosis (CAD) systems are demanding to capture abnormal prostate lesion patterns, available from each of the MRI sequences, to support the analysis of suspicious lesions. This modelling is nevertheless very challenging because the textural variability of lesions. Additionally, there is not a prior knowledge of which MP-MRI sequence could better support lesion characterization, or how these sequences should be integrated. In such sense, a set of approaches have been proposed to support detection, localization and classification of regions related to prostate cancer. For instance, Chan *et.al.* [15] proposed a multimodal approach that integrated ADC maps, and proton density (PD) images, concatenating only peripheral zone regions, which thereafter were mapped to a Support Vector Machine (SVM). Likewise, Langer *et. al* [16] integrated ADC and T2-Weighted maps into a logistic regression classifier to predict potential regions associated with cancer. The strategy works under a linear combination criteria, which could be a strong constraint. Then, Mehrtash *et.*

al [17] proposed a 3D multimodal-convolutional network that received zonal information of the prostatic gland and input regions from ADC, Maximum B Value (B-VAL), and K^{trans} sequences. Despite this approach describes lesions from an integrated early fusion strategy, there is not a clear contribution of each MRI sequence, and how the fusion strategy allows a multi-parametric characterization of malignant lesions. Afterwards, Liu *et. al* proposed a MP-MRI convolutional network that uses K^{trans} , ADC and T2WI regions to characterize suspicious relevant lesions related to prostate cancer[18]. Nevertheless, the early fusion strategy proposed in this study considers only one slice per modality, ignoring important volumetric information of neighbor tissues, where the tumors may spread and grow. Similarly, Chen *et. al* characterized malignant prostate lesions in regions of T2WI, ADC and K^{trans} images using an early fusion strategy on the InceptionV3 architecture [19]. Nevertheless, the analysis of each MRI sequence is limited by including only one slice, losing important spatial information of prostate lesions. Sunoqrot *et. al*[20] characterized clinically significant regions using a logistic regression over T2WI sequences, but ignoring antrofibromuscular-stroma regions of the prostate gland. Afterward, Bleker *et. al* [21] extracted radiomic features but losing only considering lesions that belong to the peripheral zone.

This work introduces a new multimodal deep strategy to characterize clinically significant regions, that are associated with prostate cancer, over different MP-MRI sequences. A main contribution of this approach is the multimodal inception module that allows a non-linear integration of each MP-MRI sequence, preserving spatial information of the tissues and operating on a high level of representation. The proposed approach starts by extracting potential 3D regions of interest (ROIs) around expert annotations over different MRI sequences. These regions were computed from peripheral (PZ), transitional (TZ) and anterior fibromuscular stroma (AS) zones of the prostate. Then, a set of convolutional learned filters are applied to each of the MRI sequences to extract textural patterns over micro-circulation, morphological, and cellular density features. At intermediate stages, these textural patterns are fused by using an inception-based scheme to learn cross-correlation patterns among different MP-MRI sequences. This strategy allows the integration of different prostate cancer biomarkers through a joint representation that integrates textural features from MP-MRI sequences that correlated with the disease.

2. MP-MRI sequences to support lesion prediction

MP-MRI is a fundamental tool to support cancer disease diagnosis, where different sequences capture textural parameters related to the anatomy, micro-circulation and cellular density features of the prostatic gland [8, 12]. Actually, some clinical studies have demonstrated the importance of such diagnostic sequences, even independently, to localize tumors at different prostate zones [22, 23, 24]. Also, these sequences have been used effectively to conclude in non-specific cases with negative biopsies and highly positive PSA blood test [8]. Likewise, these sequences help to diagnose and characterize prostate lesions, located far from the rectal wall, which could not be studied through a digital rectal examination (DRE) or trans-rectal ultrasound guided biopsy [8]. Specifically, from MP-MRI sequences it is possible to obtain modalities

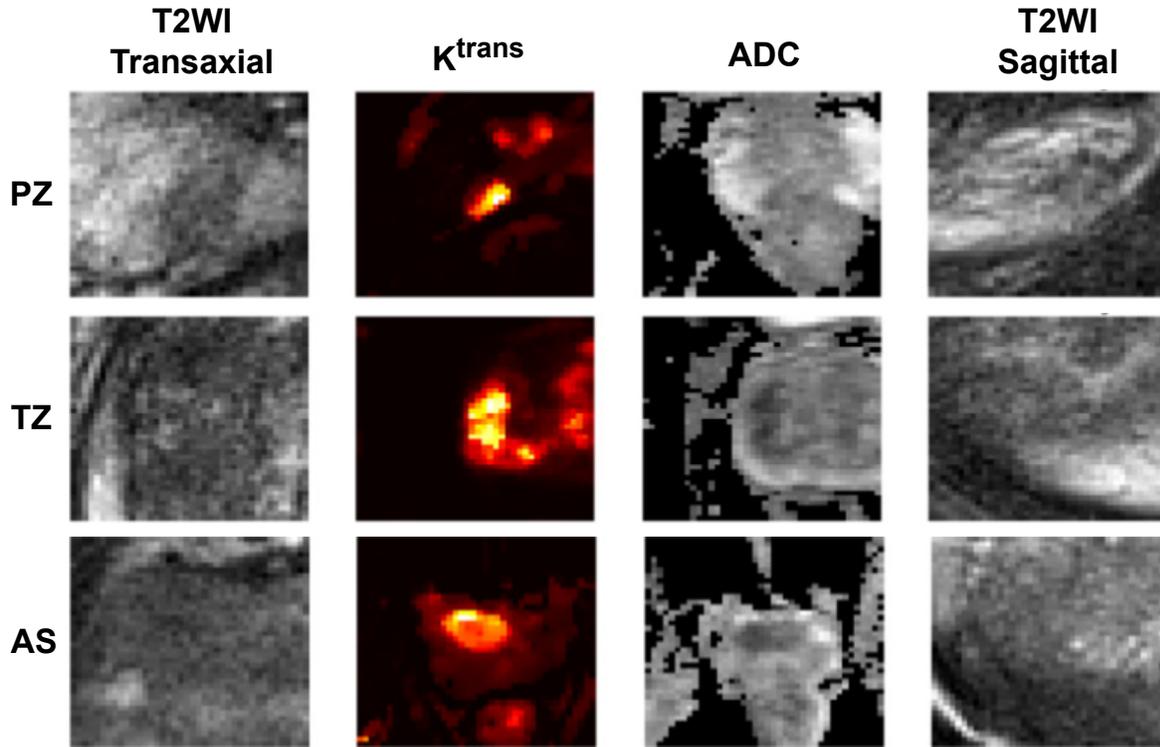


Figure 1. Prostate lesion regions taken from clinical annotations made by radiologist over different prostate zones, and projected over different parameter sequences. These delimited regions were identified as clinically significant prostatic lesions confirmed by biopsy. Each of the columns corresponds to a different sequence, from left to right the figure illustrates: T2WI-MRI trans-axial plane, K^{trans} , ADC, and T2WI-MRI sagittal plane. From top to bottom the figure shows the peripheral (PZ), transition (TZ), and antro-fibromuscular stroma (AS) zones.

(sequences) available from different capture settings, which spatially allow to identify different features of the prostatic tissue. In fact, some clinical protocols, such as PI-RADS (Prostate Imaging Reporting and Data System), recommends to use a multi-parameter observational approach integrating almost three modalities to localize and diagnose prostate cancer [25]. The most common MP-MRI sequences used in clinical routine are briefly introduced in next subsections.

2.1. Dynamic Contrast Enhanced (DCE) and K^{trans} images

Dynamic Contrast Enhanced (DCE) sequences allow to measure and localize the accumulation of contrast agents such as gadolinium in the prostatic tissue [26]. From these sequences, the measure of capillary permeability for each voxel is arranged on special K^{trans} images, which have recently emerged as an alternative to characterize and track the aggressiveness of malignant tumors. These sequences, according to *European Society of Urogenital Radiology* (ESUR), allow to observe vascular and micro-circulation properties of the tissue, among others, the plasma blood flow, vascular permeability, and the surface capillary area by unity of mass [12]. These properties fully correlates with the non controlled formation of blood vessels

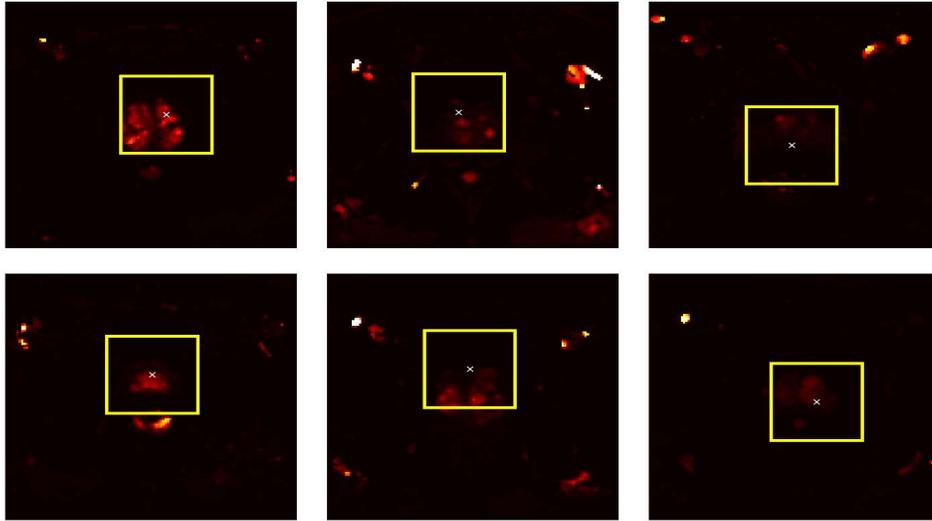


Figure 2. K^{trans} samples: from left to right it is represented the transition (TZ), antro fibromuscular stroma (AS) and peripheral zone (PZ) respectively. The row at the top represents benign prostate lesions and the bottom row illustrates positive cases of prostate cancer.

(angiogenesis), which is an essential process in the propagation of tumors in tissues [27]. Interestingly enough, in some studies have been reported a correlation between K^{trans} images and the histopatologic grade of gliomas, which could result fundamental to determine cancer degree from macroscopical observations [28, 29]. Despite of this remarked points, the sensibility on DCE-MRI is affected with some observational evidences of Angiogenesis, which could be associated to a natural process of wound healing [7]. In clinical routine is recommended to complement an integrated observational study and evaluation of prostate lesion from different MP-MRI sequences.

2.2. DWI and ADC maps

The DWI (Diffusion Weighted Imaging) represents the cellular density of the tissues as a magnitude of diffusion water particles. These sequences commonly support the discrimination between cancerous tissues and regular tissues. In such modality, the high cellular density is expressed as a low reflected intensity signal, and a non-uniform gradient change around typical uniform regions, due to the high cellular density present in cancerous tissue. Some studies have supported that the ADC (Apparent Diffusion Coefficient) maps have a negative correlation w.r.t the Gleason Grade score measured over corresponding histopathological examples [30]. Nevertheless, a main limitation of these sequences is the poor resolution that difficults the proper localization of prostate lesions in clinical routine.

2.3. T2WI Sequences

The T2 weighted image (T2WI) is one of the most used MP-MRI sequence due the anatomical description of the prostate to detect and identify lesions from multiple planes (transaxial, coronal and sagittal planes) [8]. This sequence results from relaxation time of several tissues

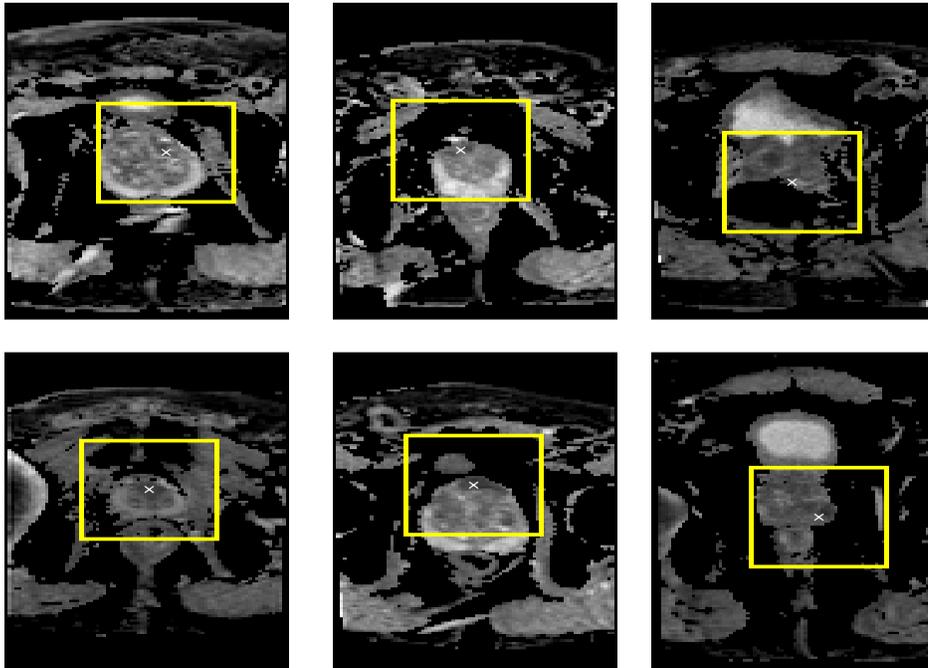


Figure 3. ADC samples: from left to right it is represented the transition (TZ), antro fibromuscular stroma (AS) and peripheral zones (PZ) respectively. The row at the top shows benign prostatic lesions and the row at the bottom illustrates positive cases of prostate cancer.

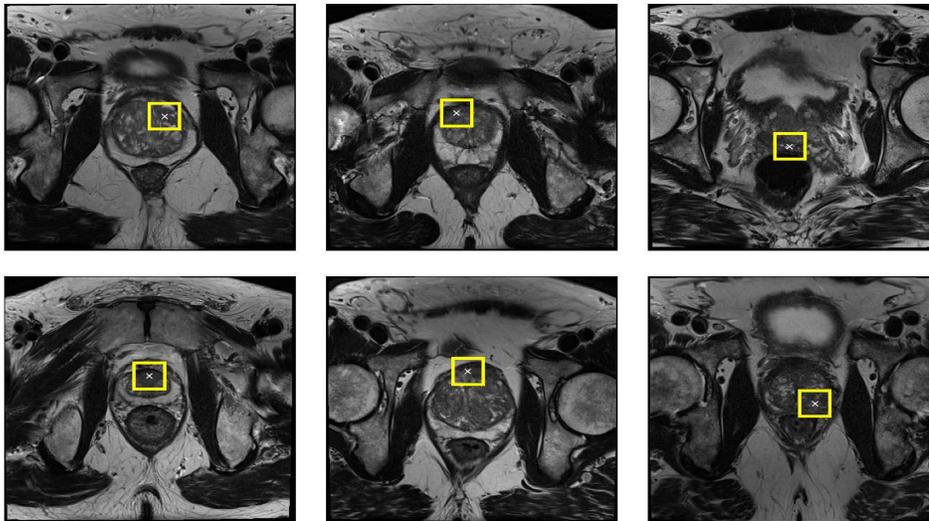


Figure 4. T2WI samples of trans-axial plane: From left to right it is reflected the transition (TZ), antro fibromuscular stroma (AS) and peripheral zone (PZ) respectively, The row at the top represents benign prostate lesions and the bottom row illustrates positive cases of prostate cancer.

and the water response. Particularly, for prostatic observations, the peripheral zone usually presents a high intensity due to the water levels, while cancerous tissues show low intensity levels. However, these intensity levels at T2WI sequences could present a high variability and some clinical analysis could be easily misdiagnosed with other pathologies such as prostatitis,

benign prostate hyperplasia (BPH) and hemorrhage post-biopsy [31].

3. Proposed approach

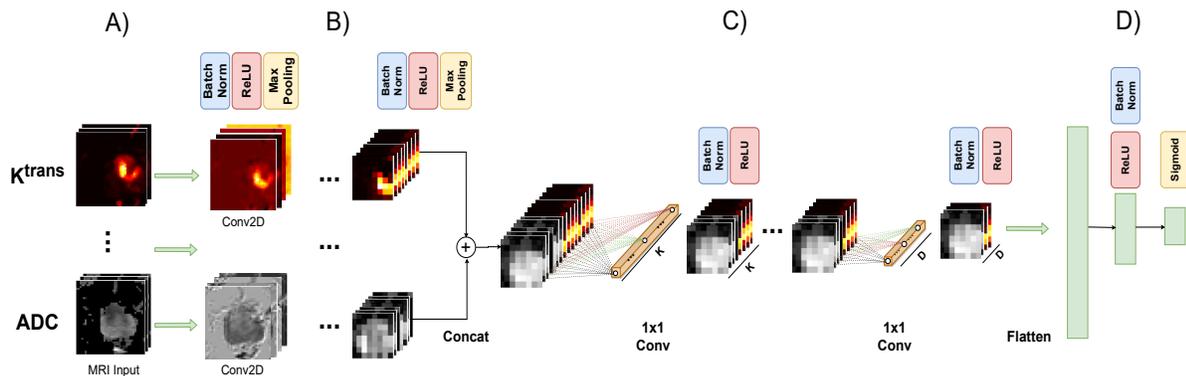


Figure 5. The general framework of our proposed strategy. In section A, different MP-MRI sequences are extracted to characterize malignant lesions. Then in section B, each sequence is decomposed in a set of convolutional filters that represents the lesions. Therefore in section C, these representations are fused using a compact cross-correlation module. Finally in section D, the fused representation is flattened (vectorized representation $\mathbb{R}^{W \times H \times L} \rightarrow \mathbb{R}^{W \cdot H \cdot K}$) to an embedding vector and mapped to a probability distribution.

This work introduces a multimodal convolutional strategy that estimates the probability that a given MRI prostate lesion would be clinically significant. Firstly, each MP-MRI modality is modelled in independent convolutional branches to achieve a better spatial representation of observed textural patterns. Each deep modality branch is then integrated with Inception-like modules, which obtain a non-linear joint representation of prostate lesions in MP-MRI. This Inception multimodal integration allows to preserve spatial information while naturally exploits the integration among activations of multiple modalities. The pipeline of our proposed strategy is illustrated in Figure 5.

3.1. Convolutional stage

In the clinical routine, an MRI study of a patient is formed by n possible MRI sequences $\{M_1, M_2, \dots, M_k, \dots, M_n\}$ with a set of prostate lesion coordinates provided by radiologists $P_i : \{\vec{p}_1, \vec{p}_2, \dots, \vec{p}_k, \dots, \vec{p}_n\}$. Then, around this localizations, we define prostate lesion as a set of cropped volumetric regions $C_i : \{c_1, c_2, \dots, c_k, \dots, c_n\}$, as shown in Figure 6.

Here, the deep representation, initially consider n isolated convolutional branches w.r.t each of the available MRI sequences. Each convolutional branch k codes information as a set of stack transformations of separated band responses of frequency filters with some coverage of mid frequencies. These responses are progressively convolved to obtain more localized and selective high-level patterns on successive convolutional layers, augmenting invariance representation. At the end of each convolutional branch, it is obtained a visual representation that captures the most relevant concepts related to the disease. At each branch, it

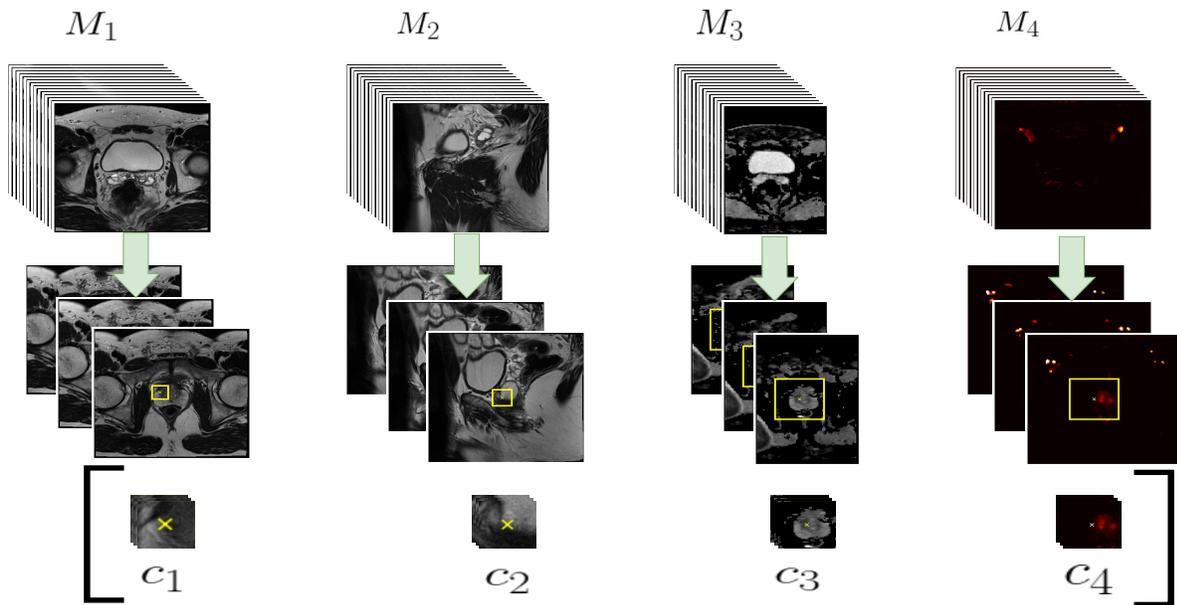


Figure 6. An Example of the MRI region extraction process with 3 slices and a window size of 40 voxels. Resulting in a ROI $C_k \in \mathbf{R}^{40 \times 40 \times 3}$.

was implemented a batch normalization over the slices axis to center the mean and reduce the variability of spatial deep features. The process of convolution for each modality is illustrated in Figure 5-b.

3.2. MP-MRI Inception fusion stage

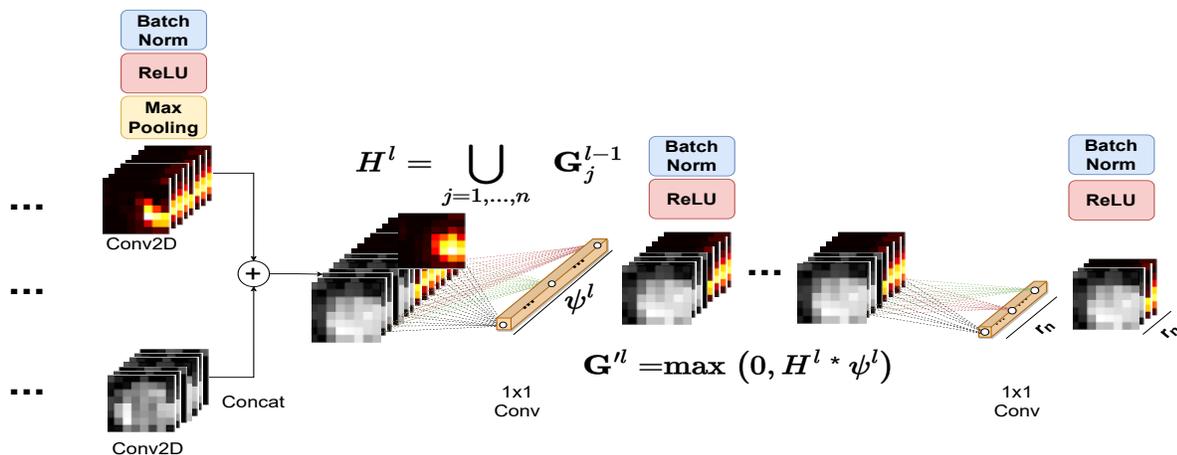


Figure 7. Module of integration of the proposed strategy, 1x1 convolutions allow our model to calculate the non-linear cross correlations between MRI ROI sequences. Then, the resulted feature map is convolved progressively with a smaller number of filters to produce a compact multimodal representation of malignant prostate lesions.

A main problem in MRI is the joint analysis of these sequences. As explained on section 2, each of these sequences illustrates different textural patterns that could be potentially used to

identify clinically significant prostate lesions. However, learning a joint representation of MRI images is a challenging task due to high variability in the textural characterization of different biologic phenomena, and the variability of the spatial resolution in MRI images. In this work, we propose a strategy to learn a joint representation of clinically significant prostate lesions in MRI images. As shown in Figure 7, for the layer l we perform a concatenation of the $l - 1$ filter responses denoted as $H^l = \bigcup_{j=1, \dots, n} \mathbf{G}_j^{l-1}$. Then, this stack of feature maps is convolved with a set of r_1 filters of size $1 \times 1 \times r_1$ denoted as: $\psi^l : \{\psi_1^l, \psi_2^l, \dots, \psi_{r_1}^l\}$. Therefore, the joint learned representation of these multimodal MRI features would be $\mathbf{G}^l = \max(0, H^l * \psi^l)$.

Such a joint learned representation \mathbf{G}^l has a non-linear integration of different MRI sequences, which allows the model to cross-correlate local features of morphological, cellular density, and blood vessel descriptors of prostate lesions over the filter space. In a similar way, the resulting fused feature map \mathbf{G}^l is convolved progressively with a set of (r_2, r_3, \dots, r_n) filters of size 1×1 , where $r_n < \dots < r_2 < r_1$. Also, at the end of these progressive blocks we apply batch normalization to reduce the variability of our multimodal MRI representation. Hence, we flatten the last layer of our integration module, following two dense layers that fully correlate multimodal hidden information, representing prostate lesion at a high semantic level.

A prostate cancer score is then achieved by mapping the embedding output to a sigmoid function to discriminate between malignant and benign prostate lesions. It should be noted that in clinical routine the most common scenario it is to have unbalanced data with respect to labeled regions. To overcome this issue, We added a penalty term between both classes as $w_0 = 1$ $w_1 = \frac{|P|+|F|}{|P|}$, where $(|P|, |F|)$ represent the cardinality of malignant (Positive to cancer) and non-malignant (False to cancer) samples during training process. Then, the loss function in such case is expressed as $L = w_0 \cdot y \cdot \log(\hat{y}) + w_1 \cdot (1 - y) \cdot \log(1 - \hat{y})$.

4. Experimental setup

4.1. Dataset

The proposed strategy was validated over the public dataset of the SPIE-AAPM-NCI Prostate MR Classification Challenge [32]. This dataset reported a set of 344 prostate MR studies captured from a retrospective study, that included T2-Weighted Imaging (T2WI), proton density-weighted imaging (PD-WI), dynamic contrast-enhanced (DCE) and diffusion weighted imaging (DWI) over the PZ, TZ and AS prostate zones. The present dataset is compounded by a training cohort of 320 annotated lesions, and a testing cohort of 208 annotated lesions. Some of the MR prostate studies may include more than a single suspicious lesion from a same patient. For such a reason, the population of lesions belongs to 204 and 140 patients for training and testing cohorts respectively. In addition, each of these MR studies included relative coordinates to the prostate lesion in each of the MRI sequences, the prostate zone of study, the clinical significance of the lesion (which is supported by biopsy), and some metadata related to the findings and patient identifiers.

The public challenge has two independent sets: training and test. Regarding training set, the evaluation of the proposed approach followed a stratified K fold cross-validation (with $K = 8$).

Additionally, for each of the $K = 8$ folds, we extracted the 20% of the training lesions under a stratified selection strategy to obtain a validation cohort. Hence, for the best representations, we submitted the lesion classification over the test set to the challenge official website and recovered the performance of the proposed approach. Finally, the performance of proposed strategy, in the classification task, was validated with respect to the Area under the receiver operating characteristic curve (ROC-AUC) because the reported in-variance to threshold values, measuring true positive rate (sensitivity) against the false positive rate (fall-out).

4.2. Architecture configuration

The proposed strategy was coded with a total of two convolutional layers for each branch that receive input modalities. Then, a total of three inception modules allowed to learn multimodal cross information through the activations. The learning scheme was implemented with an Adam optimization rule with a learning rate $lr = 1 \times 10^{-3}$. A per-voxel scaling was carried out to achieve better conditions during optimization process. Also, during the learning phase, the non-malignant lesions were penalized with the following class weights configuration: $W_0 = 1$ and $W_1 = 1 + \frac{P+F}{P}$, being P the number of samples with clinically significant lesions at and F the number of samples with non clinically significant lesions. A data augmentation was also considered, including rigid transformations, such as: flips, rotations and translations.

5. Evaluation and Results

An ablation study was firstly herein conducted to validate the proposed strategy at different MRI input configurations. For experiments with a single MRI sequence, only convolutional layers were included into scheme, together with fully connected layers. In this experiment the following MRI sequences were analyzed: the K^{trans} , ADC, and T2WI sequences. For multimodal MRI input configurations, a complete scheme was validated using the convolutional stage, fusion stage and the clinical significance prediction from dense fully connected vectors. These configurations considered the bi-modal integration of: ($K^{trans} + K^{trans}$ divergence descriptor), and ($K^{trans} + ADC$). Also, trimodal experiments were herein considered, such as: ($K^{trans} + ADC + T2WI$ transaxial) and ($K^{trans} + K^{trans}$ divergence descriptor + ADC). In the whole configuration, the proposed strategy was validated using a stratified 8 fold cross validation scheme with the annotated data of the training cohort, and regarding the test subset managed by the authors of the dataset. In Figure 8 is illustrated the general performance of different models for validation-test experiments, being remarkable the independent performance of K^{trans} and ADC sequences, achieving a ROC-AUC of 0.765 – 0.72 and 0.669 – 0.73, respectively.

Regarding multimodal experiments, the best configuration was achieved by integrating K^{trans} -ADC, achieving a ROC-AUC of 0.769 – 0.820, on validation-test experiments. Then, it was included the T2WI sequence as a third modality to complement the characterization of the disease. The results show a similar ROC-AUC performance without enhancing the representation of malignant lesions. Moreover, the inclusion of such new modality to the

proposed strategy increases the standard deviation, resulting in less confidence to the obtained results. Also, the generation of an artificial modality, divergence maps of K^{trans} , achieved stable results but comparable with the modelling of only two raw modalities. These experiments are also supported by previous works [7, 8], which have evidenced that MP-MRI enhances the characterization of clinically significant lesions on prostate cancer, combining information such as anatomical structure with blood architectural patterns. Complementary, Figure 10 illustrates the ROC-AUC mean and standard deviation reached by MP-MRI models over the PZ, AS and TZ in the validation set. As figure shows, ($K^{trans} + ADC$) model reaches the best ROC-AUC performance over the PZ (0.84). Previous studies have also shown that prostate cancers arises mainly in PZ [33, 34], which has a major tendency to illustrate cancer as vascular and cellular density abnormal patterns. In addition, it is shown that the integration of T2WI-MRI sequence at the model ($K^{trans} + ADC + T2WI$) enhances the ROC-AUC performance over the TZ zone. However, this model presents high variability in malignant prostate lesions characterization which could be explained due to homogeneous intensities between non-malignant and malignant tissues in the T2WI-MRI sequence[8]. Finally, the integration of K^{trans} divergence images at ($K^{trans} + K^{trans}$ divergence + ADC) model allow to identify malignant prostate lesions invariably to the prostate zone.

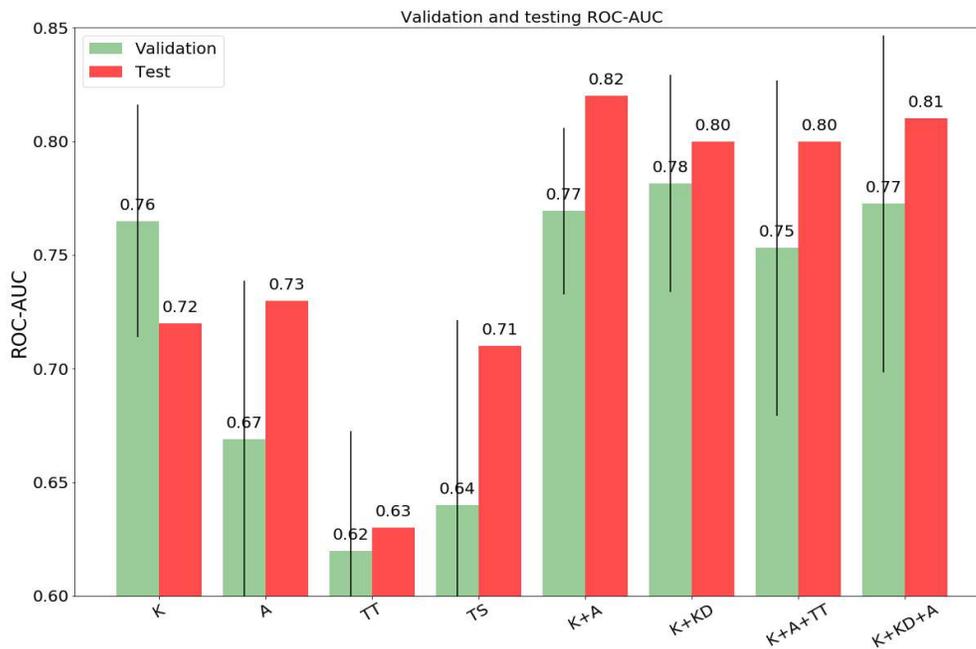


Figure 8. Results for the proposed approach, green colored bars represent the value of ROC-AUC achieved during the training stage, red colored bars represent the ROC-AUC performance of the models at the testing stage. K, KD, A, TT, TS stand for K^{trans} , K^{trans} divergence, ADC, T2WI transaxial plane and T2WI Sagittal plane respectively.

The experiments herein reported shows that a model that integrates K^{trans} images with

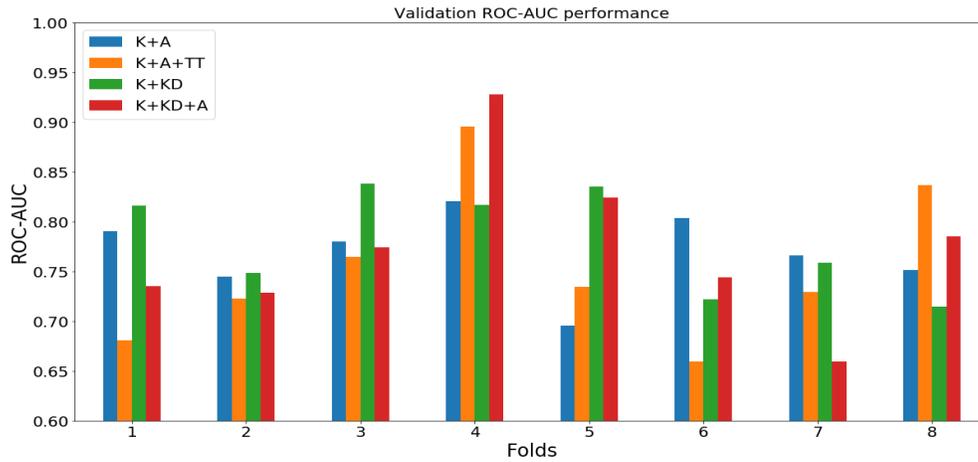


Figure 9. Validation over the training stage. An 8-fold cross validation strategy was applied over the training samples to compare the ROC-AUC performance of different multimodal models at each fold of data. Blue colored bars represent the performance of a bimodal model that integrates K^{trans} images with K^{trans} divergence images. Orange colored bars represent the performance of a bimodal strategy that integrates K^{trans} images with ADC sequences. Green colored bars represent the performance of a trimodal approach that integrates K^{trans} images, K^{trans} divergence images and ADC. Finally, red colored bars represent the performance of a CNN that integrates K^{trans} images, ADC sequences and T2WI transaxial plane sequences.

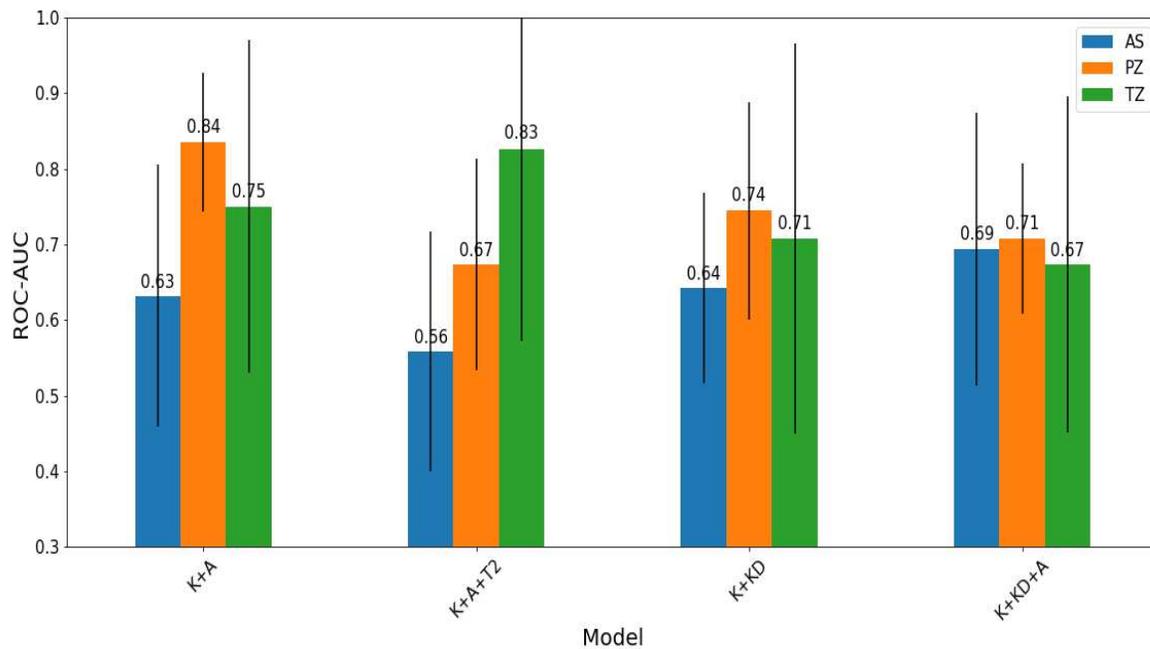


Figure 10. MP-MRI models performance in the validation set of data, blue bars represent Antrofibr muscular Stroma zones (AS), orange bars correspond to Peripheral zone (PZ) and green bars show the Transition zone (TZ)

ADC sequences achieves the best results. Hence, this model, ($K^{trans} + \text{ADC}$), was selected to analyze its performance w.r.t the window size, from axial plane, related to voxels around lesion and number of slices, along depth axis. Figure 11 summarize the obtained results, regarding ROC-AUC, for both evaluated parameters: window size (top) and number of slices (bottom). Firstly, the regions were fixed to three slices and window size was augmented from 20×20 to 80×80 . Best results were achieved by small regions, specifically, with (40×40) and (20) with a ROC-AUC on test of 0.82 and 0.81, respectively. This performance could be explained with the fact that neighborhood voxels form the appearance characteristic of lesion. In contrast, larger window sizes could introduce noise variability to the lesion model. In a second evaluation, from a fixed window size of (40×40) , it was evaluated different number of slice configurations (see on Figure 11, bottom). Remarkably, the use of only three slices resulted as ideal configuration of the proposed architecture's input, achieving a ROC-AUC of 0.82. Such fact has also associated the assumption that not only close voxels represent the lesion but also could the low resolution of images in this axis (~ 14 slices in average).

Additionally, the inception module has the capability to recover spatial patterns across the modalities by using different sizes for the kernels. In this work were evaluated different configurations of kernels of (1×1) , (3×3) , (5×5) . Also mixture blocks, *e.g.* the combination of (1×1) plus the (3×3) , were herein evaluated. Nevertheless, the best configuration was achieved by using the simplest configuration, *i.e.*, only a block from a kernel of (1×1) with the score of 0.77 – 0.82 in validation-test. In such sense, it could be understood that only across modalities integration is relevant on the model, while the spatial integration (increasing the kernel size) does not improve the characterization of lesions.

Figure 12 illustrates the output activations for independent K^{trans} and ADC branches, computed over malignant (Top rows) and benignant lesions (Buttom rows). Also, the final columns are the output attention maps, computed from a classical Grad-CAM algorithm. As expected, the K^{trans} activation responses shows regions that highlight mainly the areas where the lesion is located, which usually are the zones with higher concentration of gadolinium. For benign lesions, the activations are spread-out along the region which may explain the Angiogenesis. Interestingly, the attention maps recovered our deep representation stand out and summarize such information about the lesion. On the other hand, ADC region activations illustrate a negative correlation with the signal intensities of the lesions. In same sense, GradCAM map highlights the peripheral zone of the prostate gland, indicating low cellular density in the affected zones. Therefore, according to our qualitative and qualitative results, seems that the proposed architecture is learning the cross correlation between high contrast agent areas jointly with cellular density zones represented as low signal intensities in ADC maps. These findings are supported in the state-of-the-art by some studies that have shown a negative correlation between ADC and tumor aggressiveness quantified as Gleason score [35, 36, 37].

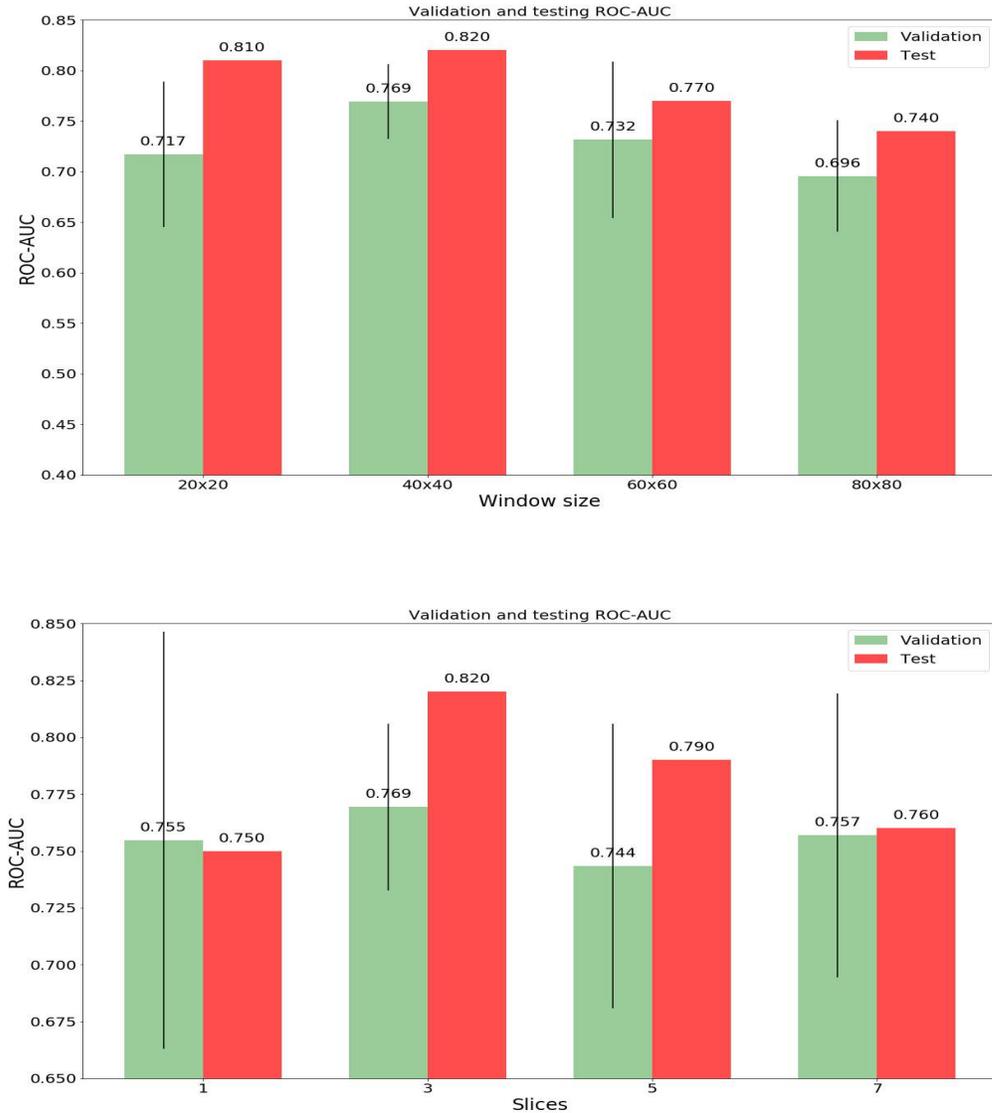


Figure 11. Panel Top: Testing and validation results of the K^{trans} ADC inspired inception strategy fixing the amount of slices and varying the window size in steps of 20 voxels. Panel Bottom: Testing and validation results of the K^{trans} ADC inspired inception strategy fixing the window size and varying the amount of slices.

6. Discussion

This work presented a MP-MRI fusion strategy that using $(1 \times 1 \times r_1)$ convolution kernels obtained cross correlations among spatial features of the available sequences. This fusion strategy obtained a ROC-AUC of 0.82 in the PROSTATEx challenge using a deep compact visual representation that integrated ADC and K^{trans} feature maps. Also, the experiments show that using global regions among trans-axial axis (similar to Mehrtash work [17]) lead to misclassifications that could be explained due to the inclusion of neighboring vital organs over

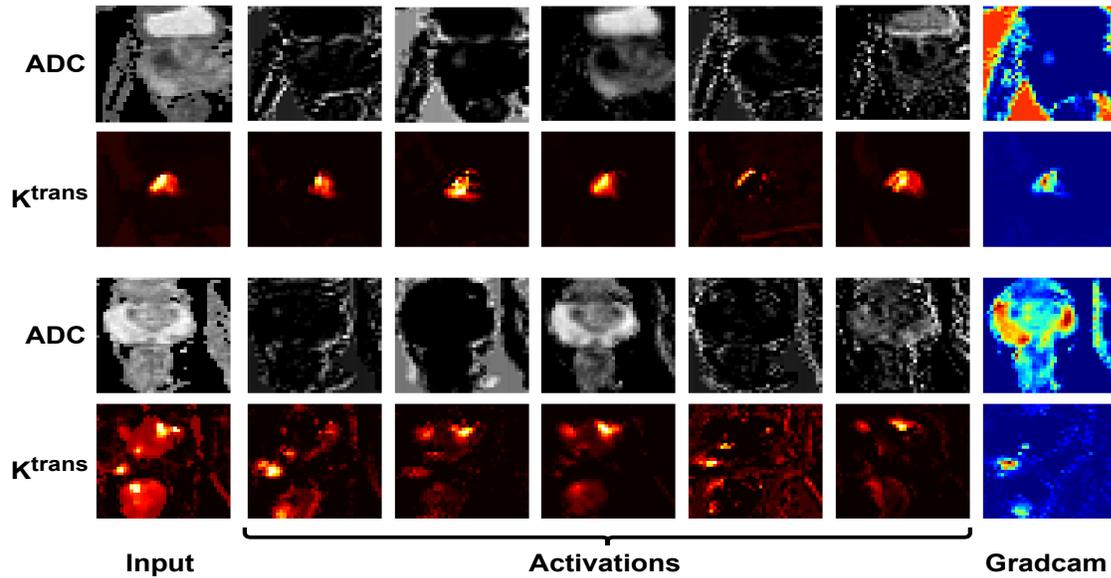


Figure 12. Visual representation obtained by the proposed architecture over a malignant (first two rows), and a benignant (last two rows) lesions. From left to right we have: the original input lesion, five deep feature maps obtained by our proposed architecture, and finally the attention map obtained by GradCAM algorithm.

MRI sequences such as the bleeder, introducing noise on the characterization of the disease. On the other hand, including very local regions (only 1 slice of the axial plane) could result on losing textural information among neighboring regions of the tissue, where the tumors could spread and grow.

To compare the performance on the diagnosis (ROC-AUC), and the computational cost of some state-of-the-art strategies (described in the quantity of parameters to be learned), a baseline comparison was carried out with similar deep architectures, evaluating their performance on the PROSTATEx challenge. As Table 1 shows, the proposed approach achieved competitive results using a compact deep architecture, which results fundamental on clinical scenarios where computational infrastructure, and the amount of labeled lesions may be limited. For instance, the best result was achieved by Liu’s approach [18] but requiring more than 4 millions of parameters to be learned. In fact, to obtain the reported results, Liu required a deep model 29 times bigger in terms of learnable parameters to obtain an improvement of only 0.02 in the ROC-AUC. Moreover, the proposed approach achieved similar results, and using only two MRI sequences (ADC + K^{trans}). Likewise, Chen used a transfer learning strategy to overcome the lack of annotated data and to pre-train his deep visual representation from the open ImageNet dataset [19]. However, Chen’s approach achieved similar results with respect to our approach but using a considerable deeper architecture, indicating that such deep representation may be limited to the main textural differences among natural and medical imaging domains. In general, these models require more computational resources to fit and adjust their representation.

The proposed compact deep visual representation has similar results compared with state

Study	ROC-AUC	Number of learned parameters
Liu’s approach [18]	0.84	4.523.298
Chen’s approach TL-ImageNet [19]	0.83	> 14, 714, 688
Our Approach	0.82	151.630
Mehrtash’s approach [17]	0.80	270.330 + 6.144 F

Table 1. Computational cost expressed as number of learnable parameters, ROC-AUC performance of different state-of-the-art studies in prostate cancer.

of the art studies due to the relationship between available data and amount of learnable parameters. From this approach, we are able to encode MRI lesions into compact non-linear cross correlation patterns, among different MRI sequences, that represent malignant lesions from a multi-parametric perspective. Specifically, in Figure 12, from the attention maps recovered was observed the complementary role between ADC and K^{trans} images. In such case, K^{trans} is dedicated to localize lesions, while the ADC show an inverse relation regarding the aggressiveness of the lesion. We hypothesize that compact models have major chance to be implemented in clinical scenarios, without specialized infrastructure. These approaches may perform fast inferences, acting online during the expert radiologist analysis. With such considerations, the model could be also updated with new samples generated in production environments, to better deal with visual variability of lesions. Finally, an important limitation of this work is the non-inclusion of T2WI MRI sequences to analyze prostate cancer regions. Nevertheless, the main role of T2WI in prostate cancer is to localize tumors, which not necessarily are malignant. In this sense, seasoned radiologists could manually annotate potential cancerous regions, and the presented approach could be introduced as a tool to support the diagnosis of prostate cancer by characterizing malignant suspicious tumors over complementary K^{trans} and ADC maps.

7. Conclusions and Perspectives

This work presented a MP-MRI deep convolutional strategy to estimate non-linear textural patterns related to prostate cancer disease. The results suggest that ADC and K^{trans} are the most relevant MRI sequences to characterize clinically significant prostate lesions under the multimodal-fusion proposed scheme. The proposed strategy boosts the ROC-AUC performance about a 10% with respect to baseline strategies. We believe that such gain in performance could be attributed to the 1×1 convolutions that learned non-linear cross correlations between textural patches of MRI sequences. These kind of convolutions allowed to reduce the dimensionality over the filter space to design a compact multimodal representation of malignant prostate lesions related to the disease. Future work will include the use of self-supervised learning strategies to pre-train deep learning models without any human annotations, reducing the effort of physicians and radiologists during annotation stages.

Acknowledgements

The authors acknowledge the Vicerrectoría de Investigación y Extensión (VIE) of the Universidad Industrial de Santander for supporting this research work by the project: “Predicción de patologías cardíacas utilizando representaciones de aprendizaje profundo en secuencias de resonancia magnética cardíaca (CMR), with SIVIE code 2703.

Compliance with ethical standards

This research study was conducted retrospectively using human subject data made available in open access by the 2017 SPIE Medical Imaging Symposium on the SPIE-AAPM-NCI Prostate MR Classification Challenge [32]. Ethical approval was not required as confirmed by the license attached with the open access data.

References

- [1] Siegel R L, Miller K D and Jemal A 2020 *CA: a cancer journal for clinicians* **70** 7–30
- [2] Polascik T J, Oesterling J E and Partin A W 1999 *The Journal of urology* **162** 293–306
- [3] Smith D S and Catalona W J 1995 *Urology* **45** 70–74
- [4] Greene D, Ali A, Kinsella N *et al.* 2015 Transrectal ultrasound and prostatic biopsy: Guidelines & recommendations for training
- [5] Ahmed H U, El-Shater Bosaily A, Brown L C, Kaplan R S, Colaco-Moraes Y, Ward K, Hindley R G, Freeman A, Kirkham A K, Oldroyd R *et al.* 2016 The promis study: A paired-cohort, blinded confirmatory study evaluating the accuracy of multi-parametric mri and trus biopsy in men with an elevated psa.
- [6] Shahait M, Degheili J, El-Merhi F, Tamim H and Nasr R 2016 *International braz j urol* **42** 60–68
- [7] Barentsz J O, Richenberg J, Clements R, Choyke P, Verma S, Villeirs G, Rouviere O, Logager V and Fütterer J J 2012 *European radiology* **22** 746–757
- [8] Murphy G, Haider M, Ghai S and Sreeharsha B 2013 *American Journal of Roentgenology* **201** 1229–1238
- [9] Oto A, Yang C, Kayhan A, Tretiakova M, Antic T, Schmid-Tannwald C, Eggenner S, Karczmar G S and Stadler W M 2011 *American Journal of Roentgenology* **197** 1382–1390
- [10] Woodfield C A, Tung G A, Grand D J, Pezzullo J A, Machan J T and Renzulli J F 2010 *American Journal of Roentgenology* **194** W316–W322
- [11] Berman R M *et al.* 2016 *Abdominal Radiology* **41** 844–853
- [12] Cuenod C and Balvay D 2013 *Diagnostic and interventional imaging* **94** 1187–1204
- [13] Penn J 2008 *Retinal and choroidal angiogenesis* (Springer Science & Business Media)
- [14] Li J, Zhang Y P and Kirsner R S 2003 *Microscopy research and technique* **60** 107–114
- [15] Chan I, Wells III W, Mulkern R V, Haker S, Zhang J, Zou K H, Maier S E and Tempny C M 2003 *Medical physics* **30** 2390–2398
- [16] Langer D L, Van der Kwast T H, Evans A J, Trachtenberg J, Wilson B C and Haider M A 2009 *Journal of Magnetic Resonance Imaging: An Official Journal of the International Society for Magnetic Resonance in Medicine* **30** 327–334
- [17] Mehrtash A, Sedghi A, Ghafoorian M, Taghipour M, Tempny C M, Wells III W M, Kapur T, Mousavi P, Abolmaesumi P and Fedorov A 2017 Classification of clinical significance of mri prostate findings using 3d convolutional neural networks *Medical Imaging 2017: Computer-Aided Diagnosis* vol 10134 (International Society for Optics and Photonics) p 101342A
- [18] Liu S, Zheng H, Feng Y and Li W 2017 Prostate cancer diagnosis using deep learning with 3d multiparametric mri *Medical Imaging 2017: Computer-Aided Diagnosis* vol 10134 (International Society for Optics and Photonics) p 1013428

- [19] Chen Q, Xu X, Hu S, Li X, Zou Q and Li Y 2017 A transfer learning approach for classification of clinical significant prostate cancers from mpMRI scans *Medical Imaging 2017: Computer-Aided Diagnosis* vol 10134 (International Society for Optics and Photonics) p 101344F
- [20] Sunoqrot M R, Nketiah G A, Selnæs K M, Bathen T F and Elschot M 2020 *Magnetic Resonance Materials in Physics, Biology and Medicine* 1–13
- [21] Bleker J *et al.* 2020 *European radiology* **30** 1313–1324
- [22] van As N J, de Souza N M, Riches S F, Morgan V A, Sohaib S A, Dearnaley D P and Parker C C 2009 *European urology* **56** 981–988
- [23] Miao H, Fukatsu H and Ishigaki T 2007 *European journal of radiology* **61** 297–302
- [24] Girouin N, Mège-Lechevallier F, Senes A T, Bissery A, Rabilloud M, Maréchal J M, Colombel M, Lyonnet D and Rouvière O 2007 *European radiology* **17** 1498–1509
- [25] Hamoen E H, de Rooij M, Witjes J A, Barentsz J O and Rovers M M 2015 *European urology* **67** 1112–1121
- [26] Essig M, Shiroishi M S, Nguyen T B, Saake M, Provenzale J M, Enterline D, Anzalone N, Dörfler A, Rovira À, Wintermark M *et al.* 2013 *American Journal of Roentgenology* **200** 24–34
- [27] Verma S, Turkbey B, Muradyan N, Rajesh A, Cornud F, Haider M A, Choyke P L and Harisinghani M 2012 *American Journal of Roentgenology* **198** 1277–1288
- [28] Blake M A and Kalra M K 2008 *Imaging in oncology* vol 143 (Springer Science & Business Media)
- [29] Zhang N, Zhang L, Qiu B, Meng L, Wang X and Hou B L 2012 *Journal of Magnetic Resonance Imaging* **36** 355–363
- [30] Hambrock T, Somford D M, Huisman H J, van Oort I M, Witjes J A, Hulsbergen-van de Kaa C A, Scheenen T and Barentsz J O 2011 *Radiology* **259** 453–461
- [31] Thompson J, Lawrentschuk N, Frydenberg M, Thompson L, Stricker P and Usanz 2013 *BJU international* **112** 6–20
- [32] Litjens G, Debats O, Barentsz J, Karssemeijer N and Huisman H 2014 *IEEE transactions on medical imaging* **33** 1083–1092
- [33] Grignon D J and Sakr W A 1994 *Journal of cellular biochemistry. Supplement* **19** 267–269
- [34] Lee J J, Thomas I C, Nolley R, Ferrari M, Brooks J D and Leppert J T 2015 *The Prostate* **75** 183–190
- [35] Gibbs P, Liney G P, Pickles M D, Zelhof B, Rodrigues G and Turnbull L W 2009 *Investigative radiology* **44** 572–576
- [36] Itou Y, Nakanishi K, Narumi Y, Nishizawa Y and Tsukuma H 2011 *Journal of Magnetic Resonance Imaging* **33** 167–172
- [37] Turkbey B, Shah V P, Pang Y, Bernardo M, Xu S, Kruecker J, Locklin J, Baccala Jr A A, Rastinehad A R, Merino M J *et al.* 2011 *Radiology* **258** 488–495

**ACTA DE SUSTENTACIÓN
TRABAJO DE INVESTIGACIÓN DE MAESTRÍA EN INGENIERÍA DE SISTEMAS E
INFORMÁTICA.**

De acuerdo con el artículo 109 del Reglamento General de Posgrado (Acuerdo del Consejo Superior No 075 de 2013), los suscritos miembros del jurado del Trabajo de Investigación de Maestría "Prostate lesions characterization in MRI sequences using a deep contrastive learning framework" presentado por el estudiante de Maestría en Ingeniería de Sistemas e Informática **Yesid Alfonso Gutiérrez Guate**, código 2208454 una vez evaluada y realizada la defensa oral dan concepto de:

APROBADO NO APROBADO APLAZADO

Se envía comunicación escrita de la presente decisión tanto al estudiante como a su director a los 13 días del mes de julio de 2023.


Prof. **FABIO MARTÍNEZ CARRILLO**
Director del Trabajo de Investigación

(EISI-UIS).


Prof. **JOHN EDILSON AREVALO OVALLE**
Codirector del Trabajo de Investigación

(Broad Institute en MIT-Harvard).


Prof. **OSCAR ACOSTA TAMAYO**
Evaluador.

(Universidad de Rennes I_Francia).



Prof. **CLAUDIO DELRIEUX**
Evaluador

(Universidad Nacional del Sur-Argentina).