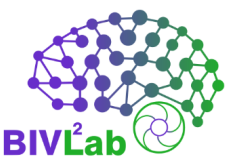
PROSTATE LESIONS CHARACTERIZATION IN MRI SEQUENCES USING A DEEP CONTRASTIVE LEARNING FRAMEWORK

YESID ALFONSO GUTIERREZ GUATE

UNIVERSIDAD INDUSTRIAL DE SANTANDER
FACULTAD DE INGENIERÍAS FISICOMECÁNICAS
ESCUELA DE INGENIERÍA DE SISTEMAS E INFORMÁTICA
BUCARAMANGA
2023



PROSTATE LESIONS CHARACTERIZATION IN MRI SEQUENCES USING A DEEP CONTRASTIVE LEARNING FRAMEWORK

YESID ALFONSO GUTIERREZ GUATE

Research work in partial fulfillment of the requirements for the degree of:
Magíster en Ingeniería de Sistemas e Informática

Advisor:
Fabio Martínez Carrillo
Ph.D. in Systems and Computer Engineering
Co-Advisor:
John Edilson Arévalo Ovalle
Ph.D. in Systems and Computer Engineering

UNIVERSIDAD INDUSTRIAL DE SANTANDER
FACULTAD DE INGENIERÍAS FISICOMECÁNICAS
ESCUELA DE INGENIERÍA DE SISTEMAS E INFORMÁTICA
BUCARAMANGA
2023

ACKNOWLEDGEMENTS

The author expresses his acknowledgment to:

The research group Biomedical Imaging, Vision and Learning Laboratory (BIVL^2ab) and each of its members. The continuous discussion between our works, background experiences, and seminars contributed very important ideas for the development of the present work. Also, I would like to make a personal acknowledgment to Professor Fabio Martínez, his personal support through a very difficult moment in my life, and his personal coaching motivated me through the whole process of this work. Moreover, he shared with me not only his scientific background but also his personal experiences to lead me through such a difficult situation. Thanks a lot professor, without your help and motivation, this research work wouldn't be possible.

In addition, I would like to thank our Faculty department of "Facultad de Ingenierías fisicomecánicas" and especially professor Olga Patricia Chacón for their economic support to attend the EMBC 2022 conference.

Additionally, I would like to thank my mother Beatríz Guate, my grandfather Alfonso Guate, and my fiancee Liseth Paola Salas, you mean all to me at this point in my life. I have no words to describe all the experiences that we have shared during our lives. Last but not least, I would like to thank my grandmother Carmen Rosa Albarracín and my uncle Carlos Alberto Guate for all the support, days, and experiences that we shared while you were alive. This research work is dedicated to you. As a family, we are going to miss you both so much in each part of our lives.

Contents

List of Figures

pág.

List of Tables

pág.

LIST OF APPENDICES

pág.

ABSTRACT

Title: Prostate lesion characterization in MRI sequences using a deep contrastive learning framework [*]

Author: Yesid Alfonso Gutiérrez Guate [**]

Keywords: prostate cancer, constrastive learning, MRI, multimodal learning

Description: Early prostate cancer diagnosis from bi-parametric MRI studies (T2WI and DWI sequences) constitutes the new guidelines in PI-RADS-2 protocol. From such sequences, malignant lesions are characterized by morphological and cellular density properties. Nonetheless, such characterization is sensible to high variability among different MRI sequences and prostate zones, which often results in misdiagnosis. Current deep learning representations have shown promising results to support the diagnosis, discriminating malignant lesions from multimodal MRI radiological findings. Nevertheless, such strategies typically require a huge amount of annotated MRI findings in multimodal MRI sequences, which clearly limits the implementation and application of these computational strategies in the clinical routine. Moreover, the learned representations may be sensible to noise generated during the data augmentation techniques. This work introduces a weakly supervised learning approach from a deep BP-MRI representation to classify malignant lesions, overcoming deep learning approaches that use multiparametric MRI. Firstly, redundant and rich tissue patches are taken from the prostate gland, allowing to adjust a representation to discriminate between lesions and healthy tissue. This pretext task is performed under a contrastive learning scheme, learning an embedding projection that groups similar patches while maximizing the distance among different classes. Then, from such representation, it is carried out a fine-tuning process to discriminate between benign and malignant lesions related to prostate cancer lesions. The proposed approach outperformed baseline studies in a public dataset, achieving a ROC-AUC of 0.85 using the 80% of the available annotated lesions. Also, using 20% of the lesions, the proposed strategy achieved a ROC-AUC of 0.80, being a promising result to transfer models to the clinical routine.

---

[*] Research work

[**] Facultad de Ingenierías Fisicomecánicas. Escuela de Ingeniería de Sistemas e Informática. Advisor: Fabio Martínez Carrillo, Ph.D.

RESUMEN

Título: Caracterización de lesiones prostáticas en secuencias MRI utilizando un marco de aprendizaje contrastivo profundo [*]

Autor: Yesid Alfonso Gutiérrez Guate [**]

Palabras clave: cáncer de próstata, aprendizaje contrastivo, MRI, aprendizaje multimodal.

Descripción: El diagnóstico temprano del cáncer de próstata a partir de estudios de MRI biparamétricos (secuencias T2WI y DWI) constituye las nuevas directrices del protocolo PI-RADS-2. A partir de dichas secuencias, las lesiones malignas se caracterizan por sus propiedades morfológicas y de densidad celular. Sin embargo, dicha caracterización es sensible a la alta variabilidad entre diferentes secuencias MRI y zonas de próstata, lo que a menudo resulta en diagnósticos erróneos. Las representaciones actuales de aprendizaje profundo han mostrado resultados prometedores para apoyar el diagnóstico. Sin embargo, estas estrategias suelen requerir una enorme cantidad de hallazgos anotados en secuencias MRI, lo que limita claramente la implementación y aplicación de estas estrategias computacionales en la rutina clínica. Además, las representaciones aprendidas pueden ser sensibles al ruido generado durante las técnicas de aumento de datos. Este trabajo introduce un enfoque de aprendizaje débilmente supervisado a partir de una representación profunda BP-MRI para clasificar lesiones malignas. En primer lugar, se toman parches de tejido redundante de la glándula prostática, lo que permite ajustar una representación para discriminar entre lesiones y tejido control. Esta tarea se realiza bajo un esquema de aprendizaje contrastivo, aprendiendo una proyección de embebidos que agrupa parches similares maximizando la distancia entre las diferentes clases. A continuación, a partir de dicha representación, se lleva a cabo un proceso de fine-tunning para discriminar entre lesiones benignas y malignas relacionadas con lesiones de cáncer de próstata. El enfoque propuesto superó a los estudios de referencia en un conjunto de datos públicos, alcanzando un ROC-AUC de 0,85 utilizando el 80% de las lesiones anotadas disponibles. Además, utilizando el 20% de las lesiones, la estrategia propuesta alcanzó un ROC-AUC de 0.80, siendo un resultado prometedor para transferir modelos a la rutina clínica.

---

[*] Trabajo de investigación

[**] Facultad de Ingenierías Fisicomecánicas. Escuela de Ingeniería de Sistemas e Informática. Director: Fabio Martínez Carrillo, Ph.D.

INTRODUCTION

The World Health Organization reported around of five million prostate cancer cases world-wide, and more than 375.000 deaths during the 2020, being the cancer with most prevalence[1]. Typically, in clinical routine, prostate cancer screening includes a Prostatic Specific Antigen (PSA), Digital Rectal Examination (DRE) and Trans-Rectal Ultrasound guided Biopsy (TRUS). Nevertheless, these current screening tests usually lead to over-diagnosis, and over-treatment, which results in a silent spread of malignant tumors[2]. For instance, the PSA reports a significant low specificity (around of 25%) related to misdiagnosis associated with other pathologies such as Prostatis, and Benign Prostate Hyperplasia (BPH), while some positive cancer cases report low PSA values[3] [4] [5]. To overcome such limitations, the PSA is carried out with DRE to confirm the disease[6] [7]. Nonetheless, the DRE is invasive and highly subjective (only 66 % of physicians can palpate correctly the

---

[1] R. L. Siegel et al. "Cancer statistics, 2020". In: *CA: a cancer journal for clinicians* 70.1 (2013), pp. 7–30.

[2] Stacy Loeb et al. "Overdiagnosis and overtreatment of prostate cancer". In: *European urology* 65.6 (2014), pp. 1046–1055.

[3] Michael J Barry. "Prostate-specific–antigen testing for early diagnosis of prostate cancer". In: *New England Journal of Medicine* 344.18 (2001), pp. 1373–1377.

[4] Benny Holmström et al. "Prostate specific antigen for early detection of prostate cancer: longitudinal study". In: *Bmj* 339 (2009).

[5] Julius Gudmundsson et al. "Genome-wide associations for benign prostatic hyperplasia reveal a genetic correlation with serum levels of PSA". in: *Nature communications* 9.1 (2018), pp. 1–8.

[6] Barry, "Prostate-specific–antigen testing for early diagnosis of prostate cancer".

[7] E David Crawford et al. "Serum prostate-specific antigen and digital rectal examination for early detection of prostate cancer in a national community-based program". In: *Urology* 47.6 (1996), pp. 863–869.

prostate gland), reporting a low agreement (kappa = 0.25)[8]. Besides, the TRUS guided biopsy is performed to analyze from a microscopical perspective the presence of carcinoma regions, but more than 30% of clinically significant cancer regions are lost[9] [10]. Moreover, there are associated risks during procedure, such as rectal bleeding and sepsis[11].

Hence, the development of a non-invasive mechanism for early diagnosis is key to develop effective treatments, which may impact the mortality index[12]. Nowadays, the multiparametric MRI sequences (MP-MRI) are a promising alternative to enhance the diagnosis, treatment, and surveillance of prostate cancer, evidencing a detection enhancement of cancer lesions, effective tumor's volume characterization[13] [14] [15], and even an in-vivo estimation of tumors progression[16]. Particularly, the T2 weighted imaing (T2WI) is used to

---

[8] Angela Zhang, Thomas Fear, and Hammood Ahmed. "Digital rectal examination in prostate cancer screening". In: *University of Western Ontario Medical Journal* 82.1 (2013), pp. 10–11.

[9] Ege Can Serefoglu et al. "How reliable is 12-core prostate biopsy procedure in the detection of prostate cancer?" In: *Canadian Urological Association Journal* 7.5-6 (2013), E293.

[10] Hashim Uddin Ahmed et al. *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.*. 2016.

[11] Mark R Quinlan, Damien Bolton, and Rowan G Casey. "The management of rectal bleeding following transrectal prostate biopsy: A review of the current literature". In: *Canadian Urological Association Journal* 12.3 (2018), E146.

[12] Siegel et al., "Cancer statistics, 2020".

[13] G. Murphy et. al. "The expanding role of MRI in prostate cancer". In: *American Journal of Roentgenology* 201.6 (2013), pp. 1229–1238.

[14] Jelle O Barentsz et al. "ESUR prostate MR guidelines 2012". In: *European radiology* 22.4 (2012), pp. 746–757.

[15] Maarten de Rooij et al. "Accuracy of multiparametric MRI for prostate cancer detection: a meta-analysis". In: *American Journal of Roentgenology* 202.2 (2014), pp. 343–351.

[16] Murphy, "The expanding role of MRI in prostate cancer".

depict the morphological features of the prostatic gland, allowing the identification of abnormal masses, or potential lesions. Complementary, diffusion-weighted imaging (DWI) allows to quantify the cellular density of the tissues as a measure of impedance in the signal, being the Apparent Diffusion Coefficient (ADC), and the Maximum B value (B-VAL), two important maps estimated from such sequences. Besides, the K^{trans} maps from dynamic contrast-enhanced (DCE) sequences allow characterization of potentially malignant lesions from micro-circulation properties, reflecting vascular patterns as a response of gadolinium influx[17] [18].

The PI-RADS (Prostate Imaging Reporting and Data System) is today the standard system to quantify, stratify and diagnose cancer lesions, involving MP-MRI sequences, allowing to standardize the interpretation, reporting, and scoring about particular suspicious lesions[19] [20]. Typically, the classical PI-RADS involves the integrated analysis of T2-Weighted Imaging (T2WI), Diffusion Weighted Imaging (DWI) and Dynamic Contrast Enhanced (DCE) sequences. Nevertheless, in the current system PI-RADS-v2 is only taken into account a bi-parametric perspective that includes T2WI and DWI sequences, avoiding the use of contrast agents, minimizing time of acquisition, associated costs and achieving

---

[17] Murphy, "The expanding role of MRI in prostate cancer".

[18] CA Cuenod and D. Balvay. "Perfusion and vascular permeability: basic concepts and measurement in DCE-CT and DCE-MRI". in: *Diagnostic and interventional imaging* 94.12 (2013), pp. 1187–1204.

[19] Barentsz et al., "ESUR prostate MR guidelines 2012".

[20] HA Vargas et al. "Updated prostate imaging reporting and data system (PIRADS v2) recommendations for the detection of clinically significant prostate cancer using multiparametric MRI: critical evaluation using whole-mount pathology as standard of reference". In: *European radiology* 26.6 (2016), pp. 1606–1612.

a relative similar diagnosis[21] [22]. Despite of the reported advances from MRI characterization, the analysis of these sequences is still fully dependent from the radiologists' experience, introducing an inter-reader variability. In fact some studies report a very moderate agreement to characterize lesion malignancy levels over the prostatic gland ($k = 0.419$), and a low agreement among radiologists to perform such interpretation over the transitional zone (TZ) ($k = 0.250$)[23].

Deep learning strategies have recently emerged as a support diagnosis tool to characterize malignant lesions, observed from MRI sequences[24] [25]. These strategies typically involve convolutional architectures, integrate different sequences, and also include other clinical variables to classify localized lesions in the assessment of prostate cancer[26]. Nonetheless, these approaches are trained from supervised schemes using the cross entropy loss function and require a huge amount of labeled and stratified data to

---

[21] Kristin K Porter et al. "Financial implications of biparametric prostate MRI". in: *Prostate cancer and prostatic diseases* 23.1 (2020), pp. 88–93.

[22] Vargas et al., "Updated prostate imaging reporting and data system (PIRADS v2) recommendations for the detection of clinically significant prostate cancer using multiparametric MRI: critical evaluation using whole-mount pathology as standard of reference".

[23] Matthew D Greer et al. "Interreader variability of prostate imaging reporting and data system version 2 in detecting and assessing prostate cancer lesions at prostate MRI". in: *AJR. American journal of roentgenology* (2019), p. 1.

[24] Alireza Mehrtash et al. "Classification of clinical significance of MRI prostate findings using 3D convolutional neural networks". In: *Medical Imaging 2017: Computer-Aided Diagnosis*. Vol. 10134. International Society for Optics and Photonics. 2017, 101342A.

[25] Quan Chen et al. "A transfer learning approach for classification of clinical significant prostate cancers from mpMRI scans". In: *Medical Imaging 2017: Computer-Aided Diagnosis*. Vol. 10134. International Society for Optics and Photonics. 2017, 101344F; Jeroen Bleker et al. "Multiparametric MRI and auto-fixed volume of interest-based radiomics signature for clinically significant peripheral zone prostate cancer". In: *European radiology* 30.3 (2020), pp. 1313–1324.

[26] Bleker et al., "Multiparametric MRI and auto-fixed volume of interest-based radiomics signature for clinically significant peripheral zone prostate cancer".

properly adjust the deep representations. Additionally, training schemes based on the cross-entropy loss function follow an inter-class minimization without considering the high variability among samples of the same class. Besides, such annotations are biased from radiologist annotations or may be labeled from biopsy output results, which results expensive, limiting the collection of representative samples to adjust deep representations.

This work we introduce a Weakly Supervised Contrastive Learning (WSCL) strategy that takes advantage of non-labeled MRI prostate regions, to achieve a significant set of training observations. The prostate gland is divided into control and lesion-affected regions, with the main purpose to adjust a deep representation that could properly discriminate lesion-affected regions from a set of weakly labeled MRI observations. Then an easy-positive mining strategy was herein implemented to find tuples of training sample configurations that better update the deep representation. Hence, from a contrastive learning scheme, a lesion affected and a batch of control regions are projected into low-dimensional embedding vectors to measure the textural similarities among control prostatic tissues, while estimating the textural differences among potentially affected zones in MRI sequences. This contrastive representation fully exploits positive and negative visual samples, by including an energy-based learning scheme that models the inter and intra-variability of prostate lesions represented as textural similarities among CSR. The proposed learning strategy was validated in three sub-sampling schemes to emulate different challenging clinical scenarios. Also, validation with supervised contrastive learning was herein included. The achieved results show that the proposed SCL strategy obtained a better performance even in smaller sub-datasets. Hence, the proposed representation is able to learn inter and intra-class variability, exploiting textural similarities among the annotated data.

1. FUNDAMENTALS

1.1. Prostate cancer diagnosis and MP-MRI sequences

During the clinical routine, the diagnosis of prostate cancer usually begins with a specific blood test namely Prostatic Specific Antigen (PSA). This diagnostic test measures the level of a glycoprotein that is present in the men's blood[27]. Nevertheless, PSA has a low specificity (approximately 25%) producing several amounts of false positives in the diagnosis of prostate cancer and requiring additional tests to obtain an accurate diagnosis. A second diagnosis test is the Digital Rectal Examination (DRE), which consists of a glove finger insertion to feel the prostate for lumpy, hard, or abnormal areas that could be related to the disease. However, the characterization of malignant lesions in DRE is limited to the physician's experience, introducing an inherent expert variability. In advanced stages, the trans-rectal ultrasound-guided biopsy is used to analyze disease from the extraction of portions of the tissue, to analyze such samples on a laboratory[28]. Nevertheless, the Trans-rectal ultrasound-guided biopsy has reported about 30% of false negatives and also, there are some studies that support the presence of secondary effects such as rectal and trans-urethral bleeding, bacteriuria and sepsis[29] [30].

Multiparametric Magnetic Resonance Imaging (MP-MRI) is a fundamental tool to support

---

[27] Barry, "Prostate-specific–antigen testing for early diagnosis of prostate cancer".

[28] D Greene, A Ali, N Kinsella, et al. *Transrectal Ultrasound and Prostatic Biopsy: Guidelines & Recommendations for Training*. 2015.

[29] Ahmed et al., *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.*.

[30] Mohammed Shahait et al. "Incidence of sepsis following transrectal ultrasound guided prostate biopsy at a tertiary-care medical center in Lebanon". In: *International braz j urol* 42.1 (2016), pp. 60–68.

the diagnosis of cancer disease, where different sequences capture textural parameters related to the anatomy, micro-circulation, and cellular density features of the prostatic gland.[31] [32]. Some clinical studies have demonstrated the importance of such diagnostic sequences to detect and localize tumors at different prostate zones[33] [34] [35]. These sequences have been used effectively to conclude in non-specific cases that result negatively in biopsy and highly positive in the PSA blood test[36]. Likewise, these sequences help to diagnose and characterize prostate lesions far from the rectal wall, which could not be studied through a digital rectal examination (DRE) or trans-rectal ultrasound guided biopsy. Specifically, from MP-MRI sequences it is possible to obtain modalities (sequences) available from different capture settings, which spatially allow the identification of different features of the prostatic tissue. In fact, some clinical protocols, such as PI-RADS (Prostate Imagining Reporting and Data System), recommend using a multi-parametric observational approach, integrating almost three modalities to localize and diagnose prostate cancer[37]. The most common MP-MRI sequences used in clinical routines are briefly explained

---

[31] Murphy, "The expanding role of MRI in prostate cancer".

[32] Cuenod and Balvay, "Perfusion and vascular permeability: basic concepts and measurement in DCE-CT and DCE-MRI".

[33] Nicholas J van As et al. "A study of diffusion-weighted magnetic resonance imaging in men with untreated localised prostate cancer on active surveillance". In: *European urology* 56.6 (2009), pp. 981–988.

[34] Huadong Miao, Hiroshi Fukatsu, and Takeo Ishigaki. "Prostate cancer detection with 3-T MRI: comparison of diffusion-weighted and T2-weighted imaging". In: *European journal of radiology* 61.2 (2007), pp. 297–302.

[35] Nicolas Girouin et al. "Prostate dynamic contrast-enhanced MRI with simple visual diagnostic criteria: is it reasonable?" In: *European radiology* 17.6 (2007), pp. 1498–1509.

[36] Murphy, "The expanding role of MRI in prostate cancer".

[37] Esther HJ Hamoen et al. "Use of the prostate imaging reporting and data system (PI-RADS) for prostate cancer detection with multiparametric magnetic resonance imaging: a diagnostic meta-analysis". In: *European urology* 67.6 (2015), pp. 1112–1121.

in the next subsections.

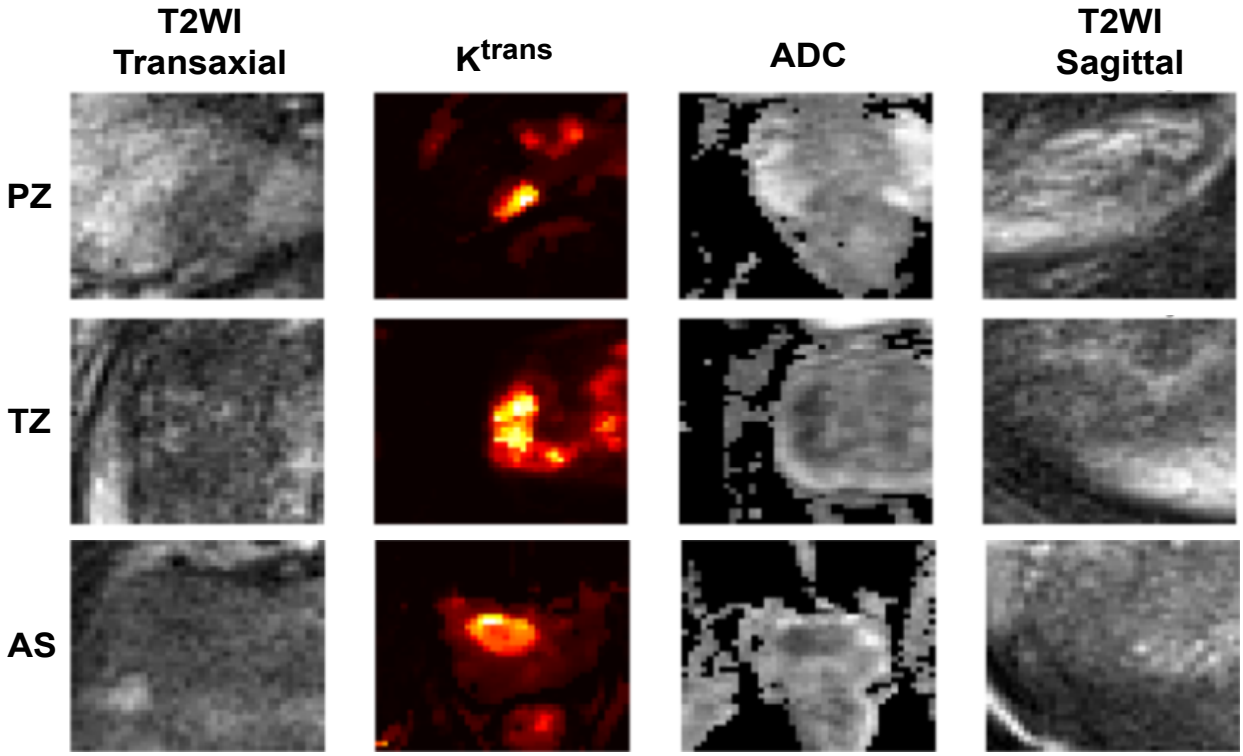


Figure 1. Clinical annotations made by radiologists over different prostate zones. These delimited regions were identified as clinically significant prostatic lesions confirmed by biopsy. Each of the columns corresponds to a different sequence, from left to right the figure illustrates: T2WI-MRI trans-axial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

1.1.1. Dynamic Contrast Enhanced (DCE) and K^{trans} images measure and localize the accumulation of contrast agents such as gadolinium in the prostatic tissue[38]. From these sequences, the measure of capillary permeability for each voxel is arranged on special K^{trans} images, which have recently evolved as an alternative to characterize and track the aggressiveness of malignant tumors. These sequences, according to ESUR (*European Society of Urogenital Radiology*), allow to observe of vascular and micro-circulation prop-

[38] Marco Essig et al. "Perfusion MRI: the five most frequently asked technical questions". In: *American Journal of Roentgenology* 200.1 (2013), pp. 24–34.

erties of the tissue, such as the plasma blood flow, vascular permeability, and the capillary surface area by a unit of mass[39].

These properties fully correlate with the non-controlled formation of blood vessels (angiogenesis), which is an essential process in the propagation of tumors in tissues[40]. Interestingly enough, some studies have reported a correlation between K^{trans} images and the histopathologic grade of gliomas. This correlation is fundamental in determining cancer degree from microscopical observations[41] [42]. Nevertheless, the sensibility of DCE-MRI is affected by some observational evidence of Angiogenesis, which could be associated with a natural process of wound healing, limiting the support of these sequences in the diagnosis[43]. In such a sense, in clinical routine is recommended to complement an integrated observational study and evaluation of prostate lesions from different MP-MRI sequences.

1.1.2. Diffusion Weighted imaging (DWI) and ADC maps come from a contrast generation method, based on the differences in the Brownian movement of diffusion input signals. Usually, this sequence allows the evaluation of the molecular function of the human body, quantified through the Apparent Diffusion Coefficient (ADC) maps. These coefficients represent the cellular density of the tissues as a magnitude of diffusion of water particles. Also, these coefficient maps commonly support the discrimination between cancerous tis-

---

[39] Cuenod and Balvay, "Perfusion and vascular permeability: basic concepts and measurement in DCE-CT and DCE-MRI".

[40] Sadhna Verma et al. "Overview of dynamic contrast-enhanced MRI in prostate cancer diagnosis and management". In: *American Journal of Roentgenology* 198.6 (2012), pp. 1277–1288.

[41] Michael A Blake and Mannudeep K Kalra. *Imaging in oncology*. Vol. 143. Springer Science & Business Media, 2008.

[42] Na Zhang et al. "Correlation of volume transfer coefficient Ktrans with histopathologic grades of gliomas". In: *Journal of Magnetic Resonance Imaging* 36.2 (2012), pp. 355–363.

[43] Barentsz et al., "ESUR prostate MR guidelines 2012".

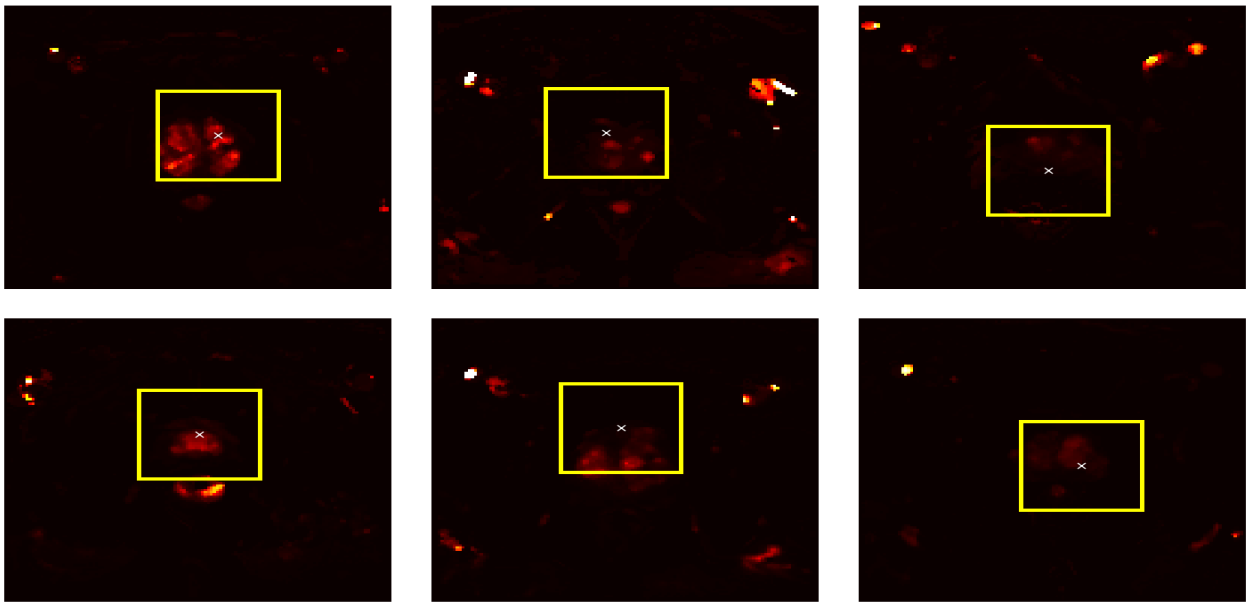


Figure 2. K^{trans} samples: from left to right it is represented the transition (TZ), antro fibromuscular stroma (AS) and peripheral zone (PZ) respectively. The row at the top represents benign prostate lesions and the bottom row illustrates positive cases of prostate cancer.

sues and regular tissues. 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 presence in cancerous tissue. Also, some studies have supported that this MRI sequence has a negative correlation w.r.t the Gleason Grade score measured over corresponding histopathological examples[44]. Nevertheless, a main limitation of this sequence is the poor resolution that difficult the proper localization of prostate lesions in clinical routine.

1.1.3. T2WI Sequences result from the relaxation time response of several tissues, taking advantage of the water differential content. This sequence allows anatomical description allowing to detect and identify the stage of prostate lesions from multiple planes (transax-

[44] Thomas Hambrock et al. "Relationship between apparent diffusion coefficients at 3.0-T MR imaging and Gleason grade in peripheral zone prostate cancer". In: *Radiology* 259.2 (2011), pp. 453–461.

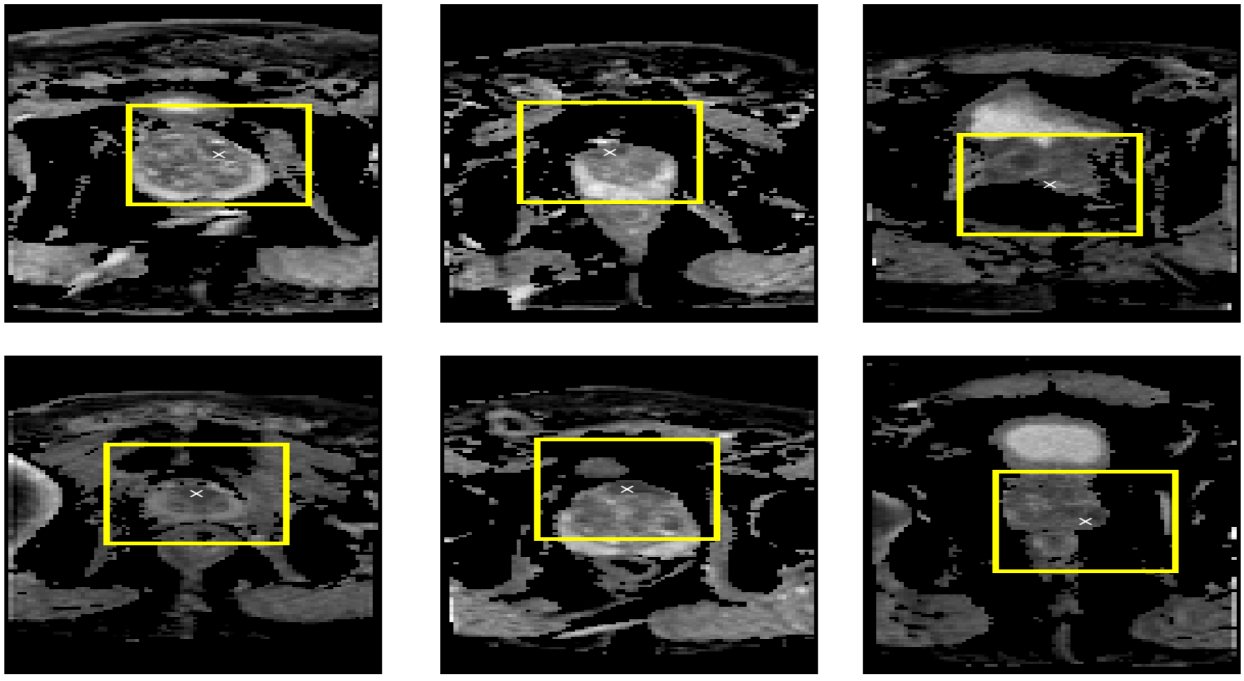


Figure 3. ADC samples: from left to right it is represented the transition (TZ), antro fibromuscular stroma (AS) and peripheral zones (PZ) respectively. The row at the top shows benign prostatic lesions and the row at the bottom illustrates positive cases of prostate cancer.

ial, coronal, and sagittal planes)[45]. Particularly, for prostatic observations, the peripheral zone usually presents high signal intensities due to the water levels in such zone, while cancerous tissues show low-intensity levels. However, these intensity levels at T2WI sequences could present a high variability depending on the zone and the properties of lesions. Therefore, some clinical analyses could be easily misdiagnosed with other pathologies such as prostatitis, benign prostate hyperplasia (BPH), and hemorrhage post-biopsy[46].

---

[45] Murphy, "The expanding role of MRI in prostate cancer".

[46] 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.

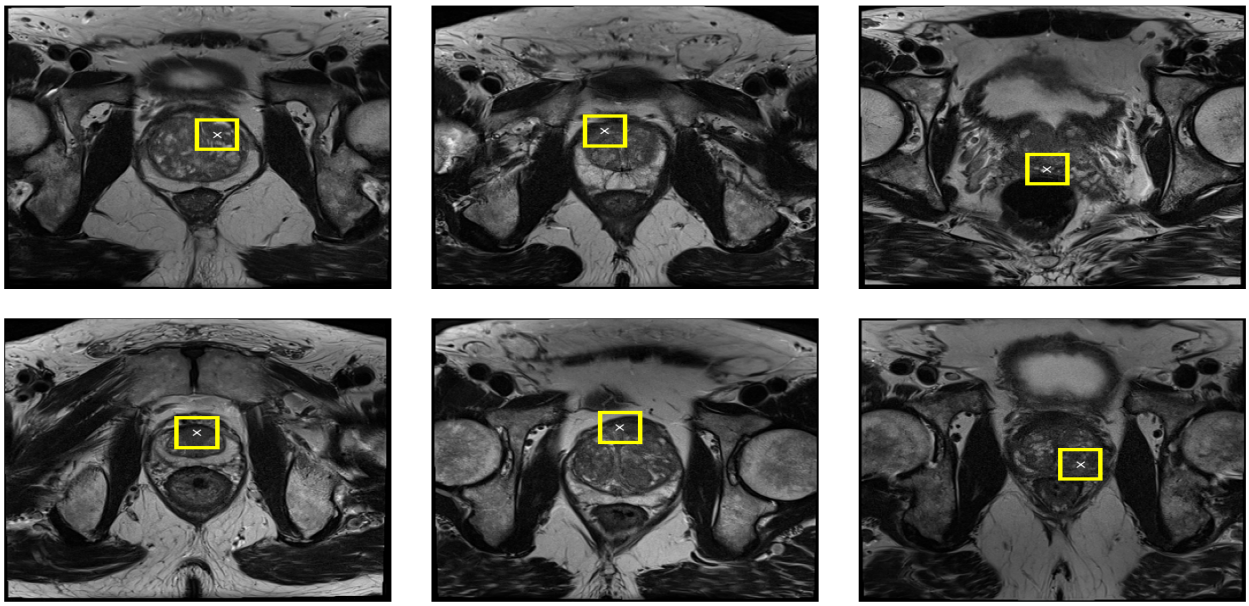


Figure 4. T2WI samples of trans-axial plane: From left to right it is reflected the transition (TZ), antro fibromuscular stroma (AS) and peripheral zone (PZ) respectively, The row at the top represents benign prostate lesions and the bottom row illustrates positive cases of prostate cancer.

1.2. Contrastive Learning

Many of the strategies to classify prostate lesions underlie on CNN representations f_θ, learning convolutional kernels to exploit visual features. This discrimination learning is achieved through a cross-entropy optimization rule that adjusts the visual representations during several training batch iterations and follows a supervised learning framework. Specifically, this learning scheme takes a set of MRI prostate lesions $x = \{x^1, x^2, \ldots, x^n\}$, and a set of clinical significance labels $y = \{y^1, y^2, \ldots, y^n\}$, to update the learnable parameters $\theta = [\theta_1, \theta_2, \ldots, \theta_m]$ of a visual representation f_θ. For doing so, the set of training MRI lesions is forward propagated through the model to obtain malignant probabilities of each lesion $\hat{y} = f_\theta(x)$. Then, the cross entropy loss function is computed to measure the error between the diagnosis predictions $\hat{y}$ and the ground truth biopsy labels y as follows:

$$\mathrm{L}(\hat{\mathrm{y}}, \mathrm{y}) = -\frac{1}{\mathrm{n}} \sum_{i=1}^{\mathrm{n}} y^i \cdot \log \hat{y^i} + (1 - y^i) \cdot \log (1 - \hat{y^i}) \tag{1}$$

The binary cross-entropy loss is optimized to minimize the error between the diagnosis and the biopsy labels with respect to the set of learnable parameters θ, and it's back-propagated through the model f_θ to obtain the best visual representation of malignant prostate lesions. This learning rule only focuses the attention on discriminative patterns that minimize error among trained classes. Nonetheless, there exist some reported limitations such as the sensibility to noisy labels that can affect the proper boundary among classes. Besides, there exist limitations related to few label data for training and statistical bias on unbalanced datasets[47] [48]. Unfortunately, in clinical scenarios, it is common to face unbalanced datasets, and in general, these datasets have relatively few data w.r.t the high variability of sample observations. Even worst, in real scenarios, the updating of the model with new samples may be corrupted by noisy labels that result from the reported experts' bias.

Contrastive learning has emerged as an alternative scheme to learn visual representations from pattern similarities among samples, with main advantages over scenarios with

---

[47] Sainbayar Sukhbaatar et al. "Training convolutional networks with noisy labels". In: *arXiv preprint arXiv:1406.2080* (2014).

[48] Zhilu Zhang and Mert R Sabuncu. "Generalized cross entropy loss for training deep neural networks with noisy labels". In: *arXiv preprint arXiv:1805.07836* (2018).

scarce of labeled data[49] [50] [51] [52]. Firstly, a pretext task is selected to obtain a deep visual representation related to the downstream (target) task. Then, the obtained deep representation is transferred to the target domain, under the hypothesis that a sufficiently general visual representation may share representation for multiple tasks. Moreover, this general representation can deal with high intra-variability and inter-visual samples. For instance, Oord et al[53] learned representations from a contrastive predictive coding scheme that allowed to estimate "future" windowed neighboring regions x_{t+k} from a context image representation c_t, and the current window image x_t. For doing so, Oord proposed to measure a dense ratio $f(x_{t+k}, c_t)$ that preserves the mutual information (MI) among the future windowed neighboring regions and the context image representation as follows: $f(x_{t+k}, c_t) = exp(z_{t+k}^T W_k c_t)$, where $z_{t+k} = g_{enc}(x_{t+k})$ are the neighboring regions x_{t+k} projected on a latent space through an encoder network g_{enc}, $c_t = g_{arr}(\{z_i\}); \forall_{i<t}$. This latent context representation uses an auto-regressive decoder network g_{arr}, and W_k which is a transformation for the prediction (a linear transformation on the author's experiments).

Then, using these features from the original image, Oord proposes (InfoNCE) a loss objective function based on the noise contrast estimation (NCE) as follows:

$$L = -\mathbf{E}_X \left[log \frac{f(x_{t+k}, c_t)}{\sum_{i,i \neq t}^{N} f(x_i, c_t)} \right]$$

---

[49] Aaron van den Oord, Yazhe Li, and Oriol Vinyals. "Representation learning with contrastive predictive coding". In: *arXiv preprint arXiv:1807.03748* (2018).

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

[51] Kaiming He et al. "Momentum contrast for unsupervised visual representation learning". In: *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. 2020, pp. 9729–9738.

[52] Prannay Khosla et al. "Supervised contrastive learning". In: *arXiv preprint arXiv:2004.11362* (2020).

[53] Oord, Li, and Vinyals, "Representation learning with contrastive predictive coding".

, where $X = \{x_k\}$ is the set of all the windowed samples, x_t is the positive sample, and $x_i; i \neq t$ are $N - 1$ negative samples. The expectation $\mathbf{E}_X$ over whole X samples define the optimization rule. In this sense, the InfoNCE loss maximizes the mutual information among signal samples that are close with respect to the time t, while minimizing the mutual information of samples that are far from time t.

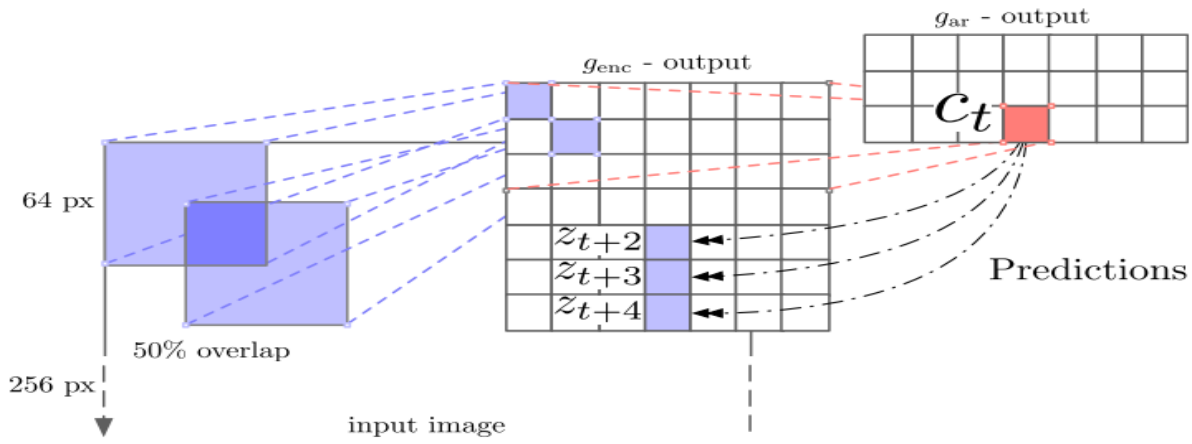


Figure 5. Oord's work: Firstly, window patches x_i are encoded to a latent vector space z_i. All of these latent vectors are mapped to a context representation c_t using all the latent vectors that belong to previously observed patches ($i < t$). Then, Oord's approach maximizes the mutual information among neighboring window patches, while minimizing the mutual information of other window patches. Finally, this method is able to estimate neighboring window patches x_{k+t} projected on the latent space z_{k+t} given a context window c_t.

In Figure 5 is shown the pipeline of Oord's work to estimate neighboring window patches z_{k+t}, from a set of encoded vectors that compound the context (c_t). The final obtained visual representation was evaluated on the open ImageNet dataset with promising results, considering the fact that there were no labeled data during the training process. In fact, a self-learning task was carried out from the context of image information taking advantage of dense visual information, and enriching the training dataset. Then, Chen et al[54], proposed a framework for Contrastive Learning of visual representations, whose pretext

[54] Chen et al., "A simple framework for contrastive learning of visual representations".

tasks consist of identifying if two given patch images $\tilde{x_i}, \tilde{x_j}$ belong to the same image.

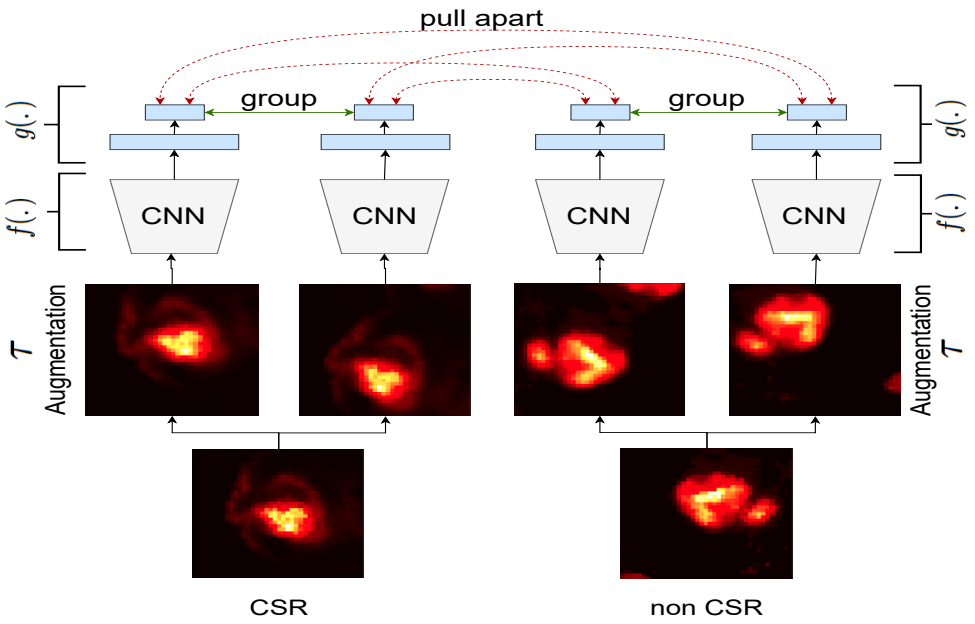


Figure 6. Adaptation of Chen's pipeline for characterization of clinically significant regions (CSR) and benign tissues (non-CSR) on K^{trans} images: The input image x is transformed to produce two versions $\tilde{x_i}, \tilde{x_j}$. Then each of these images is encoded with a CNN $f(.)$ to a high-dimensional representation h_i, h_j. Afterward, these representations are projected on a latent space using a multi-layer perceptron $g(.)$ to produce two latent vectors z_i, z_j. Finally, Chen's framework measures the similarities among these latent vectors with the purpose to maximize their agreement.

As shown in Figure 6, Chen first uses two transformation functions τ, τ' (online data augmentation) to obtain two versions $\tilde{x_i} = \tau(x), \tilde{x_j} = \tau'(x)$ of a same image x. Afterward, using a visual representation $f(.)$ (ResNet50), Chen encodes both images to a high-dimensional representation $h_i = f(\tilde{x_i}), h_j = f(\tilde{x_j})$. After that, Chen uses a Multi-Layer Perceptron (MLP) to project the representations on latent vector spaces z_i, z_j. Where $z_k = g(h_k) = W^{(2)}\sigma(W^{(1)}h_k)$, $W^{(l)}$ represents the weights of the layer l, and $\sigma(.)$ is the ReLu activation function. This process is made on all the available images to produce a set of images $\tilde{x_k}$. Each particular image has positive pairs $\tilde{x_i}, \tilde{x_j}$, and the other images $\tilde{(}x_k; k \neq i$ work as negatives. Finally, Chen's contrastive learning framework measures the

similarities among latent samples with the following expression:

$$L_{i,j} = log \frac{exp(sim(z_i, z_j)/t)}{\sum_{i=1}^{2N} \mathbf{1}_{[k \neq i]} exp(sim(z_i, z_k)/t)}$$

Where $sim(.)$ is the cosine similarity function, and t is a temperature hyper-parameter. Therefore, Chen's framework aims to learn good representations $f(.)$ that could encode these images in a latent space. The proposed approach measures the similarities among images, with respect to their directions on the latent space. Finally, Chen used the pre-trained ResNet50 $f(.)$ with a linear top classifier to transfer the learned features to the open ImageNet dataset, obtaining an accuracy of 76.5%, matching the performance of a ResNet50 under a supervised scheme. Also, He et al[55] proposed a contrastive learning strategy that models the problem of self-learning from the available data, as a dictionary lookup strategy, where given an image value x_i, the idea is to find its corresponding key k_i. In this method, He first proposed to create a dictionary $D = \{(k_i, x_i)\}$, where each key $k_i = f_k(x_i)$ is an embedding vector, which is generated from a momentum encoder network $f_k(.)$.

As observed in figure 7, He's approach considers two encoders, a query encoder $f_q(.)$ and a moment slow progressive momentum encoder $f_k(.)$. These encoders produce a set of unique identifiers tokens for the input images $\{q, k_0, k_1, \ldots, k_{n-1}\}$. Here, the pretext task identifies which is the corresponding key $\in K$ that matches or is more similar to the query q. To estimate such a similarity among these keys and queries, He used a contrastive loss function inspired by the InfoNCE as follows:

---

[55] He et al., "Momentum contrast for unsupervised visual representation learning".

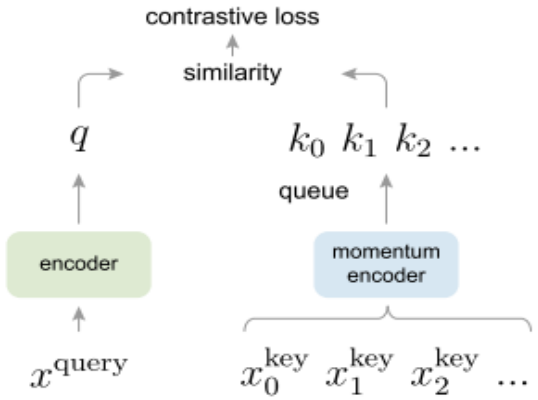


Figure 7. He's et al[56] pipeline for momentum contrastive representation learning: given an input query image x_q, and a set of key images $K = \{x_k^0, x_k^1, x_k^2, \ldots, x_k^{N-1}\}$
.

$$L_q = log \frac{exp(q.k_+/t)}{\sum_{i=0}^{N} exp(q.k_i/t)}$$

where $q = f_q(x_q)$ is the query key, $k_k \in K; k_k = f_k(x_k)$ is a key, k_+ is the positive or target key, and t is a temperature scalar. Additionally, He's approach considers an update step to the dictionary D, where a new batch of data X, will be stacked inside of the dictionary, while the oldest batch will be deleted from the dictionary to keep updating the available keys and values of the data representation. To make a slow and progressive update of the weights of the momentum encoder $f_k(.)$, it was considered an optimization rule based on the momentum $\theta_k \leftarrow m\theta_k + (1-m)\theta_q$, where m is the momentum coefficient. This learning scheme based on a dictionary lookup was evaluated on multiple downstream tasks including image classification, segmentation, and object detection. He's approach not only obtained interesting results on the ImageNet, COCO, and VOC datasets but also overcame its own supervised learning scheme version, showing that contrastive learning could be a promising strategy to pre-train deep learning models without labeled data.

Afterward, Khosla et al[57] proposed a supervised contrastive learning scheme to learn pattern similarities among samples that belong to the same class. Similarly to Oord's work, Khosla used a data augmentation module to produce two versions $\hat{x_i} = Aug(x_i)$ from the same input image x_i. Then, each patch image $\hat{x_i}$ is encoded using an encoder network $r_i = f_\theta(x_i)$ (ResNet50), to represent these patch images on high $d = 2048$-dimensional space of features $r_i \in \mathbf{R}^d$. After that, each of these embedding vectors r_i is projected on a latent space $z_i = proj(r_i)$, where $proj(.)$ is an encoding network (an MLP in this case) to map each of the high dimensional vectors to a lower latent space $z_i \in \mathbf{R}^p; p = 128$.

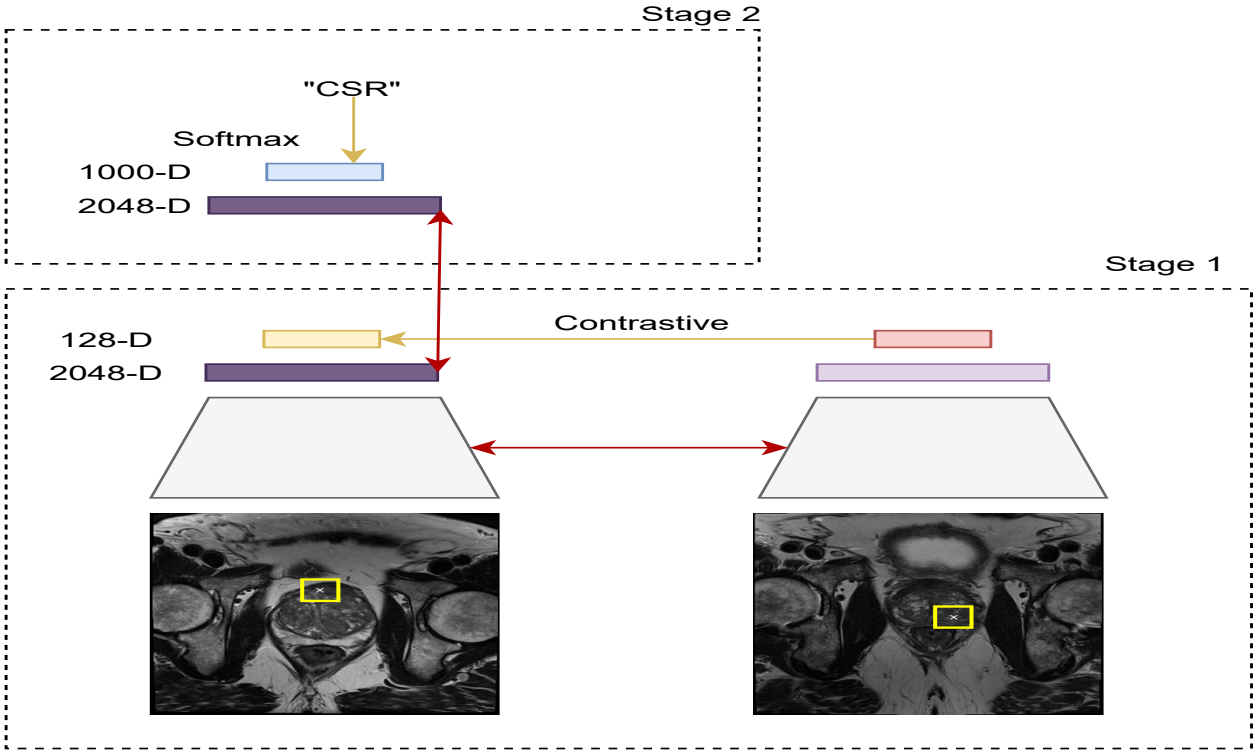


Figure 8. Adaptation of Khosla's framework for supervised contrastive learning: In this approach, the biopsy labels are used to identify if the provided T2WI sequences belong to clinically significant regions (CSR) or benign regions (non-CSR). The main idea of this framework is to learn pattern similarities among sequences of the same labeled class at stage one and to fine-tune the learned representation with a top classifier at stage two.

---

[57] Khosla et al., "Supervised contrastive learning".

The similarity among samples of the same class is performed with the supervised contrastive loss function developed by Khosla, which is expressed as:

$$L = \sum_{i=1}^{2N} \frac{-1}{|P(i)|} log\{ \sum_{p \in P(i)} \frac{exp(z_i z_p / t)}{\sum_{a \in A(i)} exp(z_i z_a / t)} \}$$

Where $A(i) = \{i\}$ is the set of anchor indexes, $P(i) = \{p \in A(i) : y_p = y_i\}$ is the set of positive indexes $p \neq i$, $|P(i)|$ is the cardinality of positives samples, and t is a temperature hyper-parameter scalar on the training scheme. This supervised contrastive learning scheme not only allowed to self-learn similarities among images that belong to the same class but also achieved outstanding results transferring learning to the ImageNet dataset, especially for the ResNet50V1 and ResNet200V1 architectures. Contrastive learning may properly learn deep representations, taking advantage of batches from positive and negative samples. This training framework may be adapted in supervised and self-supervised schemes, exploring pretext task that forces to learn, and considering more challenging scenarios. These approaches have been successfully evaluated in natural domain scenarios for multiple applications such as classification, semantic segmentation, and identification. Also, in the medical domain has been reported preliminary but potential alternatives to classification[58] [59] [60] [61].

---

[58] Khosla et al., "Supervised contrastive learning".

[59] Ting Chen et al. "Big self-supervised models are strong semi-supervised learners". In: *arXiv preprint arXiv:2006.10029* (2020).

[60] Alberto Rossi, Monica Bianchini, and Franco Scarselli. "Robust prostate cancer classification with siamese neural networks". In: *International Symposium on Visual Computing*. Springer. 2020, pp. 180–189.

[61] Ozan Ciga, Anne L Martel, and Tony Xu. "Self supervised contrastive learning for digital histopathology". In: *arXiv preprint arXiv:2011.13971* (2020).

2. COMPUTATIONAL STRATEGIES TO SUPPORT PROSTATE CANCER DIAGNOSIS

Computational strategies have allowed approximate class boundary limits to discriminate malignant prostate lesions. For instance, Chan *et. al.*[62] proposed a multimodal approach that integrated ADC maps, and proton density (PD) images to support the diagnosis of malign cancer regions. In this approach, a set of segmented regions, from different MRI sequences, were concatenated and mapped to a Support Vector Machine (SVM) to obtain an automatic classification. However, This work was carried out only in the peripheral zone, ignoring an important portion of the prostatic gland. Likewise, Langer *et. al*[63] integrated ADC and T2-Weighted maps into a logistic regression classifier to predict potential regions associated with cancer. Nevertheless, The strategy works under linear combination criteria, which could be a strong constraint in this problem. Bleker et. al[64] extracted a set of 92 radiomic features that were classified with a Random Forest (RF) algorithm, and an Extreme Gradient Boosting (XGB), but evaluated only on the peripheral zone of the prostatic gland. Also, the radiomic feature extraction approach may lose some important spatial properties of the prostate tissues, which limits the diagnosis of neighboring local regions, where tumors could grow and spread.

Currently, some end-to-end deep convolutional strategies have been proposed to inte-

---

62 Ian Chan et al. “Detection of prostate cancer by integration of line-scan diffusion, T2-mapping and T2-weighted magnetic resonance imaging; a multichannel statistical classifier”. In: *Medical physics* 30.9 (2003), pp. 2390–2398.

63 Deanna L Langer et al. “Prostate cancer detection with multi-parametric MRI: Logistic regression analysis of quantitative T2, diffusion-weighted imaging, and dynamic contrast-enhanced MRI”. in: *Journal of Magnetic Resonance Imaging: An Official Journal of the International Society for Magnetic Resonance in Medicine* 30.2 (2009), pp. 327–334.

64 Bleker et al., “Multiparametric MRI and auto-fixed volume of interest-based radiomics signature for clinically significant peripheral zone prostate cancer”.

grate multiple MRI sequences, allowing to support biopsy correlation from macroscopic findings. For instance, Giannini et al.[65] proposed a fully convolutional strategy, that from DCE, T2WI, and DWI sequences is able to select and identify malignant pixels on the prostate gland. Despite that the author aligned different MRI sequences to avoid variability in the spatial features of the tissues, the study was conducted only on the peripheral zone. This is a main drop-back because of the signal intensity variability according to the prostate zone[66]. Recently, Some approaches have proposed 3D architecture that integrated different regions of ADC, B-VAL, and K^{trans} images with zonal information of the prostatic gland[67][68]. Sunoqrot *et. all*[69] characterized clinically significant regions using a logistic regression over T2WI sequences. Despite this study developing an interesting method to normalize T2WI images autonomously from the prostate muscle and fat tissues, this study does not analyze other MP-MRI sequences that may complement the diagnosis. Similarly, Liu et al.[70] proposed the XmasNet to integrate T2WI, K^{trans} images, ADC maps, and DWI sequences. Liu's also proposed a new image data augmentation

---

65 Valentina Giannini et al. "A fully automatic computer aided diagnosis system for peripheral zone prostate cancer detection using multi-parametric magnetic resonance imaging". In: *Computerized Medical Imaging and Graphics* 46 (2015), pp. 219–226.

66 Murphy, "The expanding role of MRI in prostate cancer".

67 Mehrtash et al., "Classification of clinical significance of MRI prostate findings using 3D convolutional neural networks".

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

69 Mohammed RS Sunoqrot et al. "Automated reference tissue normalization of T2-weighted MR images of the prostate using object recognition". In: *Magnetic Resonance Materials in Physics, Biology and Medicine* (2020), pp. 1–13.

70 Saifeng Liu et al. "Prostate cancer diagnosis using deep learning with 3D multiparametric MRI". in: *Medical imaging 2017: computer-aided diagnosis*. Vol. 10134. International Society for Optics and Photonics. 2017, p. 1013428.

strategy rotating and slicing through the available 3D MRI sequences. Nevertheless, such architecture integrated only one slice per available MRI sequence, losing local neighboring tissues, where aggressive tumors could spread. Afterward, Chen et. al.[71] proposed a multimodal VGG-16 that integrated sequences of T2WI, ADC, and feature maps coded from K^{trans} images on a multimodal transfer learning scheme. Despite this strategy dealing with the scarcity of labeled prostate lesions, the textural differences between the natural and medical domains limit a deep representation.

Deep learning representations to characterize lesions from MP-MRI prostate sequences usually involve convolutional architectures with early or late fusions, or also include clinical variables[72] [73]. Such deep learning approaches deal with large observational variability by adjusting multi-level node representations in hierarchical configurations. Nonetheless, adjusting these node parameters requires a huge amount of observations to discover non-linear boundaries among classes. In fact, such labeled data should follow other conditions such as a well-balanced distribution among class samples, and confidence with respect to a proper ground truth that avoids biasing the model to wrong label samples. This last issue is a critical point in medical applications with labels that have a strong dependency on radiologist annotations with variability in the observations.

From a discriminative point of view, such models are typically trained using the cross-entropy loss function to adjust deep representations, that guide training from an inter-

---

[71] Chen et al., "A transfer learning approach for classification of clinical significant prostate cancers from mpMRI scans".

[72] Mehrtash et al., "Classification of clinical significance of MRI prostate findings using 3D convolutional neural networks".

[73] Bleker et al., "Multiparametric MRI and auto-fixed volume of interest-based radiomics signature for clinically significant peripheral zone prostate cancer".

class minimization but without additional considerations. For instance, Mehrtash et. al[74] proposed a 3D convolutional architecture that integrated different regions of ADC, B-VAL, and K^{trans} images together with zonal information of the prostate gland. However, the proposed strategy deals with the lack of labeled lesions by using an artificial data augmentation strategy, that includes translations and flippings. Nevertheless, such transformations could introduce noise in the final representation related to textural information of neighboring organs such as the bladder or the rectum. Afterward, Liu et. al proposed XMASNET, a 3D deep convolutional model inspired by the classical VGG16[75]. In this approach, Liu early integrated DWI, ADC, and K^{trans} sequences through the stacking of the sequences very similar to RGB color codification. Although the author performed an ablation study among the available sequences, the best architecture lost morphological patterns available in T2WI sequences. In addition, the proposed architecture does not discriminate lesions across prostate regions, which could introduce variability and sensitivity lost on the prediction outputs to support the diagnosis.

Subsequently, Chen et. al[76] proposed a transfer-based computational strategy from a pre-trained deep learning model to integrate T2WI sequences, ADC, and K^{trans} images. Although this strategy avoided training a relatively deep architecture, textural differences between medical and natural images limit the deep representation. Then Hung Le et. al[77] integrated T2WI and ADC sequences using two deep-learning models responsible for representing prostate lesions in a low-dimensional space. As a result, a vector encod-

---

[74] Mehrtash et al., “Classification of clinical significance of MRI prostate findings using 3D convolutional neural networks”.

[75] Liu et al., “Prostate cancer diagnosis using deep learning with 3D multiparametric MRI”.

[76] Chen et al., “A transfer learning approach for classification of clinical significant prostate cancers from mpMRI scans”.

[77] Minh Hung Le et al. “Automated diagnosis of prostate cancer in multi-parametric MRI based on multimodal convolutional neural networks”. In: *Physics in Medicine & Biology* 62.16 (2017), p. 6497.

ing was obtained to represent the textural similarity between morphological and cellular density patterns for the same tissue. Then, Tsehay et. al[78] proposed a three-channel convolution scheme to integrate T2WI, ADC, and B-Value=2000 images, that follows deep supervision to characterize prostate lesions at different scales. On the other hand, Yang et al[79] proposed a strategy that first identify regions from T2WI images, and then used a co-trained deep learning model from patterns related to cellular density (ADC), and morphology (T2WI). This approach estimated a probability map allowing to localize tumor malignancy but from a co-learning perspective, the representation may be sensitive to false-positive tumors.

More recently, Wang et al[80] proposed a weakly supervised strategy to localize prostate lesions from an end-to-end learning scheme of two staked nets: 1) one dedicated to the detection and multimodal co-registration and the other 2) dedicated to the detection of prostate lesions. Both nets are mutually trained from a weakly supervised scheme with the pretext task to identify slices with potential lesions, guiding the extraction of relevant features. The model included overlapping, consistency, and classification loss functions to find potential prostate lesion localization. This approach nonetheless may have strong dependencies on co-registration tasks with may propagate error localization to the lesioned network. Also, this model is limited to include intensity variability that represents prostate tissues at different zones. Subsequently, Bleker et. al[81] extracted radiomic features, which

---

[78] Yohannes K Tsehay et al. “Convolutional neural network based deep-learning architecture for prostate cancer detection on multiparametric magnetic resonance images”. In: *Medical imaging 2017: Computer-aided diagnosis*. Vol. 10134. SPIE. 2017, pp. 20–30.

[79] Xin Yang et al. “Co-trained convolutional neural networks for automated detection of prostate cancer in multi-parametric MRI”. in: *Medical image analysis* 42 (2017), pp. 212–227.

[80] Zhiwei Wang et al. “Automated detection of clinically significant prostate cancer in mp-MRI images based on an end-to-end deep neural network”. In: *IEEE transactions on medical imaging* 37.5 (2018), pp. 1127–1139.

[81] Bleker et al., “Multiparametric MRI and auto-fixed volume of interest-based radiomics signature for clini-

were classified with a Random Forest (RF) algorithm and the Extreme Gradient Boosting (XGB) algorithm. However, this work was performed only in the peripheral zone of the prostate gland. Additionally, Aldoj et al[82], proposed a multimodal 3D convolutional architecture using different combinations of T2WI, DWI, ADC sequences, and K^{trans} images.

Moreover in a previous work[83], we proposed a 3D multimodal late fusion learning strategy using a LeNet-based architecture that integrated K^{trans} images and T2WI MRI sequences to characterize malignant lesions. Nevertheless, our late fusion strategy lacks deep end-to-end learning, since the late fusion depends on a voting scheme with fixed α values. Finally, we also proposed an Inception-based 3D architecture[84] that integrated ADC and K^{trans} maps through 1x1 convolutional modules cross-correlating the textural patterns of these MRI maps. However, our proposed scheme does not include any rigid registration modules to deal with the alignment of the MRI sequences. Moreover, the difference in resolution and coordinates among sequences may affect the proposed architecture, including surrounding lesion regions that could introduce noise in the final representation.

---

cally significant peripheral zone prostate cancer".

[82] Aldoj et al., "Semi-automatic classification of prostate cancer on multi-parametric MR imaging using a multi-channel 3D convolutional neural network".

[83] Yesid Gutiérrez, John Arevalo, and Fabio Martıénez. "A Ktrans deep characterization to measure clinical significance regions on prostate cancer". In: *15th International Symposium on Medical Information Processing and Analysis*. Vol. 11330. SPIE. 2020, pp. 80–88.

[84] Yesid Gutiérrez, John Arevalo, and Fabio Martıénez. "An inception-based deep multiparametric net to classify clinical significance MRI regions of prostate cancer". In: *Physics in Medicine & Biology* 67.22 (2022), p. 225004.

3. RESEARCH PROBLEM

The American Cancer Society has reported more than 190.000 new prostate cancer cases during 2020, being the major incidence cancer in men[85]. The prostate screening methods used in the clinical routine are generally invasive, report a high number of false positives, and have reported the presence of some side effects[86] [87] [88]. Recently, Multi-parametric magnetic resonance imaging (MP-MRI) has emerged as a non-invasive alternative to diagnose the disease, allowing the characterization of morphological and vascular prostate features[89]. This analysis is a complex task that depends on the experience of seasoned radiologists, who must locate prostate lesions regions and also determine their clinical relevance, being a costly procedure with highly variable results regarding the biopsy[90].

Recently, deep learning representations have demonstrated remarked capabilities to deal with image variability in supervised tasks. Nonetheless, the effective implementation of this framework requires a considerable amount of training data, a main limitation on clinical scenarios. Specifically, these strategies require MRI sequences labeled with spatial location and annotated with Gleason correspondence, which results in a tedious and expensive task to be carried out in large studies. The typical cross-entropy rule has the main

---

85 Siegel et al., “Cancer statistics, 2020”.

86 Barry, “Prostate-specific–antigen testing for early diagnosis of prostate cancer”.

87 Ahmed et al., *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.*.

88 Shahait et al., “Incidence of sepsis following transrectal ultrasound guided prostate biopsy at a tertiary-care medical center in Lebanon”.

89 Barentsz et al., “ESUR prostate MR guidelines 2012”.

90 Murphy, “The expanding role of MRI in prostate cancer”.

responsibility of these requirements to minimize inter-class samples but losing information of intra-class variability. Also, forcing into a rigid supervised scheme may induce learning expert bias, which could be considered in the task of characterizing malignant lesions with respect to the biopsy gold standard. Currently, the implementation of these computational tools may be limited by missing the opportunity of learning from unbalanced datasets, and the use of alternative schemes that uses non-annotated visual information.

Research Question

How to design a contrastive learning strategy to characterize clinically significant prostate lesions over MP-MRI sequences under clinical scenarios with few labeled lesions?

4. OBJECTIVES

General Objective

- To develop a computational strategy to support the clinical significance characterization of prostate lesions using a contrastive deep learning framework.

Specific Objectives

- To select a dataset that includes two or more MRI sequences of the prostatic gland.
- To formulate a contrastive learning scheme to exploit similarities among malignant prostate tissues.
- To develop a visual representation strategy for malignant prostate lesions characterization.
- To perform a systematic evaluation of the proposed deep contrastive learning strategy simulating different clinical scenarios.

5. PROPOSED APPROACH

This work introduces a contrastive learning strategy that adjusts a deep net from a weakly labeled task that discriminates non-specific lesions from the rest of prostate volumetric tissues. Thereafter, a fine-specific adjusting was performed with a limited set of labeled samples to discriminate between benign and malignant lesions, related to prostate cancer disease. The learned embedding representation allows to automatically separate lesions diagnosed with cancer disease, from a linear class hypothesis. The complete content of this section was accepted in the Annual International Conference of the IEEE Engineering in Medicine and Biology Society - EMBC 2022[91]. Also, an extended version is under review in the Journal of Computer Methods and Programs in Biomedicine

5.1. Selection of candidate BP-MRI regions

To supply lesion label scarcity, a pretext task is firstly defined as the discrimination between lesion and control (non-lesion) prostate tissue, allowing to recover of redundant textural information, along prostate tissue, to adjust the deep convolutional scheme $f_\theta(.)$. Figure 9 summarizes the workflow to select challenging control patches that close textural information with respect to a lesion reference.

The proposed strategy extracts k multimodal MRI regions from radiological findings in the prostate gland at the lesion-affected and control regions. Here as observed in Figure 9, control samples are extracted from T2WI prostate tissue (x^k_{T2WI},), bounded from a

[91] Yesid Gutiérrez, John Arevalo, and Fabio Martínez. "Multimodal Contrastive Supervised Learning to Classify Clinical Significance MRI Regions on Prostate Cancer". In: *2022 44th Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC)*. IEEE. 2022, pp. 1682–1685.

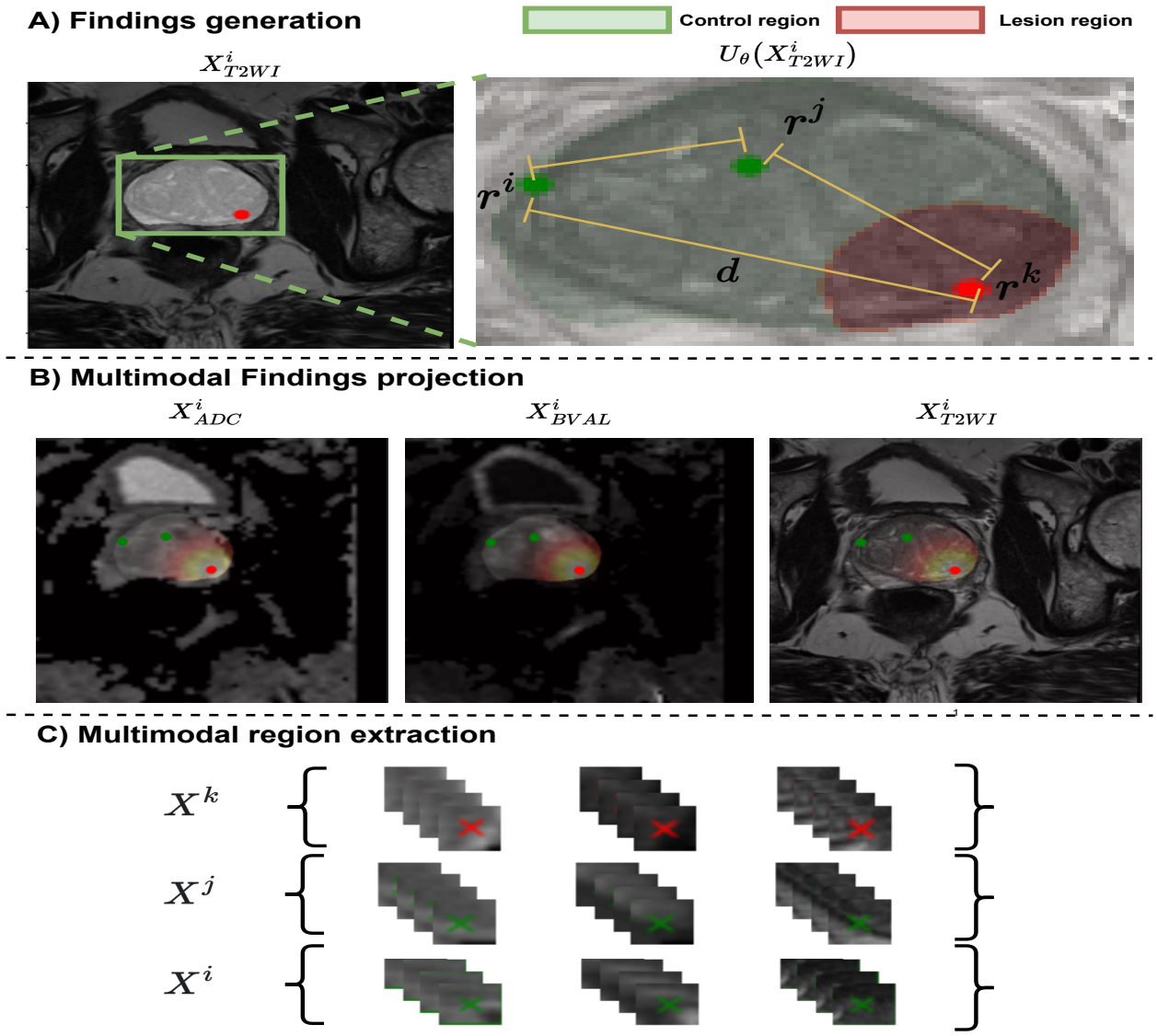


Figure 9. Extraction of candidate regions from MRI sequences. First in A), using a state-of-the-art architecture U_θ the prostate gland is delineated and divided into anatomical zones. Then, using random 2D Gaussian maps, the prostate gland is divided into control and lesion-affected regions. Afterward, in section B) these findings are projected to ADC, BVAL, and T2WI sequences in order to obtain a multimodal representation of the same regions. Finally, in section C) multimodal volumetric patches are extracted from the projected findings.

segmentation mask, $U_\theta(x^k_{T2WI})$[92].

[92] Anneke Meyer et al. “Towards patient-individual PI-Rads v2 sector map: CNN for automatic segmenta-

The lesion location samples $\{r^i, r^j, \dots r^k\}$ are taken from annotations but without considering associations to cancer disease. Afterward, over each lesion location, r^k is projected a 2D Gaussian lesion mask $G(x^k_{T2WI})$ that simulates the boundaries of the lesion-affected region. Then, from the control region $H(x^k) = U_\theta(x^k_{T2WI}) - G(x^k_{T2WI})$, M MRI candidate regions are randomly selected preserving a minimum euclidean distance d among all the selected regions. Figure 9-A illustrates the selection of prostate patches. These selected points $\{r^i, r^j, \dots r^k\}$ are then projected to each MRI sequence to carry out a bi-parametric extraction of candidate regions. In figure 9-C is shown a set of volumetric regions x^j, x^k that was extracted from the weakly labeled findings r^j, r^k for each of the BP-MRI sequences, obtaining, as a result, a set of weakly selected MRI observations related to control and lesion tissue samples.

5.2. Prostate lesions from a deep visual representation

The proposed approach is flexible to adopt any multimodal deep architecture adjusting the deep representation from a weakly supervised strategy. In this work, we implemented a state-of-the-art multimodal convolutional architecture that encodes suspicious prostate lesions into embedded vectors, integrating zonal information, and MRI sequences from independent branches[93]. In the literature, such representation was originally evaluated with input branches related to ADC, B-VAL, and K^{trans} maps, also incorporating one-hot encoding vectors with anatomical zone information. Also, the deep representation was adjusted to receive BP-MRI inputs (ADC, BVAL, T2WI), as suggested in the last PI-RADS

---

tion of prostatic zones from T2-weighted MRI". in: *2019 IEEE 16th International Symposium on Biomedical Imaging (ISBI 2019)*. IEEE. 2019, pp. 696–700.

[93] Mehrtash et al., "Classification of clinical significance of MRI prostate findings using 3D convolutional neural networks".

v2 protocol[94]. Additionally, the representation was adjusted following a typical supervised learning scheme using the cross-entropy loss function.

Contrary, we adapted the architecture to encode multimodal MRI sequences as an embedding vector following a contrastive learning scheme. Consequently, each BP-MRI region $\mathbf{x^i}$ is represented as a set of bi-parametric parameters $\mathbf{x^i} : \{x^i_{ADC}, x^i_{BVAL}, x^i_{T2WI}\}$, where $\mathbf{x^i} \in \mathbf{R}^{w \times h \times s \times p}$ is a volumetric region of $w \times h$ pixels with s slices. In such case, p is a specific MRI sequence from $\mathbf{x^i}$. Afterward, each of these volumetric parameters is mapped to independent convolutional branches $f_{\theta,p}(\mathbf{x^i_p}) \rightarrow \mathbf{z^i_p}$ that represents each sequence as a low dimensional embedding vector ($\mathbf{z^i_p} \in \mathbf{R^n}$). So, we encoded each lesion $\mathbf{x^i}$ with three independent convolutional branches $\{f^i_{\theta,p_1}, f^i_{\theta,p_2}, f^i_{\theta,p_3}\}$, obtaining as a response a set of embedding vectors $\{\mathbf{z^i_1}, \mathbf{z^i_2}, \mathbf{z^i_3}\}$. On the other hand, zonal information $\mathbf{z^i_{zone}}$ is also integrated as one hot encoding vector that represents the prostate region anatomy. Finally, the overall representation can be expressed as $z_\theta(\mathbf{x^i}) = \left[f^i_{\theta,ADC}; f^i_{\theta,BVAL}; f^i_{\theta,T2WI}; z_{zone}\right]$, obtaining concatenated embedding branches z^i_θ, where $z^i_\theta : \left[z^i_{\theta,ADC}; z^i_{\theta,BVAL}, z^i_{\theta,T2WI}; z^i_{zone}\right]$ is an embedding projection that encodes and represents BP-MRI

5.3. Contrastive Learning for Weakly-Supervised regions

Lesion characterization is here performed through a contrastive learning strategy to discriminate samples but from a geometrical embedding space. For contrastive learning, an initial step consists of the selection of batches constituted by query, positive $(\mathbf{x^i}, \mathbf{x^j})$ regions and a set of adversarial samples $\{\mathbf{x^k}\}$. The selection of these batches is key to maximizing the intra-class mutual textural information among query-positive MRI regions while maximizing the inter-class variability between query-negative regions. Hence, in this work, we adopted a lesion mining strategy that selected query-positive $(\mathbf{x^i}, \mathbf{x^j})$

---

[94] Murphy, “The expanding role of MRI in prostate cancer”; Barentsz et al., “ESUR prostate MR guidelines 2012”.

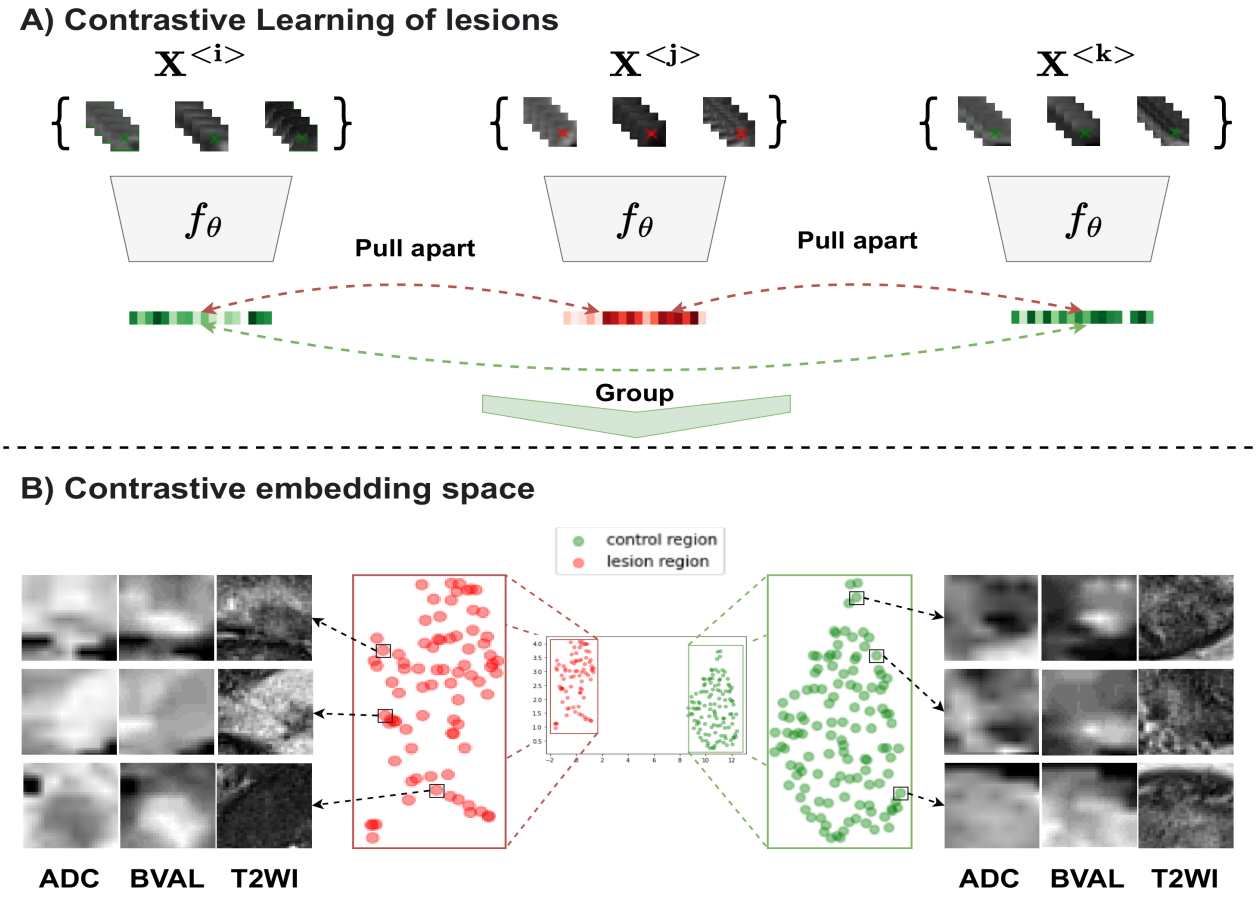


Figure 10. Weakly contrastive learning of regions. As shown in A), a set of pseudo-labeled prostate regions are projected into an embedding representation using the adjusted deep network $f_\theta(.)$. Then in B), the contrastive embedding space preserves the geometrical space, maintaining a small distance among MRI regions of the same class. Additionally, such embedding space linearly separates the classes of the MRI regions.

and query-negative $(\mathbf{x^i}, \mathbf{x^k})$ region pairs. For this purpose, we used the easy-positive mining strategy, which selects the closest positive sample $\mathbf{x^j}$ with respect to an anchor MRI region $\mathbf{x^i}$, that belongs to the same class $C(\mathbf{x^i}) = C(\mathbf{x^j})$. This selection is formally defined as $\mathbf{x^j} = \arg\min_{\mathbf{x^j}:C(\mathbf{x^j})=C(\mathbf{x^i})} d(f(\mathbf{x^i}, f(\mathbf{x^j}))$, where $d(,)$ is the euclidean distance among the encoded representation of two MRI regions. Additionally, we also included the semi-hard negative mining strategy to find the most relevant negative region $\mathbf{x^k}$, where $\mathbf{x^k} = \arg\min_{\mathbf{x^k}:C(\mathbf{x^k})\neq C(\mathbf{x^i})} d(f(\mathbf{x^i}), f(\mathbf{x^k}))$, constrained by $d(f(\mathbf{x^i}, f(\mathbf{x^k})) > d(f(\mathbf{x^i}), f(\mathbf{x^j}))$.

This mining strategy allowed us to filter the available labeled MRI data during the learning stage to obtain optimal triplets that maximize the similarity among positive regions, also preserving some textural variability with respect to negative regions[95].

Afterward, these region batches are used to train a deep architecture $f_{\hat{\theta}}$ that projects MRI regions as embedding vectors ($f_{\hat{\theta}}(\mathbf{x^i} \in \mathbf{R}^n)$, maximizing the distance among classes. Moreover, this learning scheme also constrains the embedding vector representations to remain as close as possible when they represent the same class, preserving the geometrical space with a separation between the classes (see Figure 10). To measure the textural variability and similarity of regions, we used the Normalized Temperature (NT-Xent) contrastive loss function over the embedding representation of the regions $f_{\hat{\theta}}(\mathbf{x^i})$, as follows:

$$L^{i,j} = -log\frac{exp(\left[f_{\hat{\theta}}(\mathbf{x^i})^\top \cdot f_{\hat{\theta}}(\mathbf{x^j})\right]/\tau)}{\sum_{k=1}^{2N} 1_{[k\neq i]} exp([f_{\hat{\theta}}(\mathbf{x^i})^\top \cdot f_{\hat{\theta}}(\mathbf{x^k})]/\tau)}$$

Where τ is a temperature hyper-parameter that weights the scale of the similarity measurements. Moreover, the proposed contrastive learning strategy allows to adjust and optimize the resultant embedding representation not only from the radiologists' annotations but also from the similarity measurements $L^{i,j}$ among MRI regions, which is a main advantage with respect to traditional supervised learning schemes based on the Cross-entropy loss function.

To adjust our deep representation and obtain an optimal embedding space that separates malignant and benign regions, two progressive contrastive learning stages were implemented to obtain a general representation of control and lesion-affected regions. Then, such representation is transferred and adjusted to discriminate potential malignant lesions.

---

[95] H Xuan et. al. "Improved embeddings with easy positive triplet mining". In: *Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision*. 2020, pp. 2474–2482.

In this sense during the first stage, the deep representation is updated from control and lesion regions extracted from the MRI sequences under a weekly supervised scheme. From these regions, the deep network $f_{\hat{\theta}}(.)$ is adjusted to discriminate the MRI regions from an embedding projection. Consequently, in the second stage, the deep representation is transferred from the previous stage and refined from $f_{\hat{\theta}}(.) \rightarrow f_{\theta}$ using the radiologists and biopsy annotations. During this stage, the deep model is also updated following the same contrastive learning strategy preserving the textural similarity among lesions, and the variability with respect to malignant regions related to prostate cancer disease. Thus, the proposed approach learns an optimal projection of prostate lesions into a geometrical representation achieving an optimal contrastive embedding space linearly separable. To discriminate malignant lesions a logistic regression classifier was trained to obtain a hyperplane that defines the boundary among both classes.

5.4. Experimental setup

5.4.1. Dataset This work was trained and validated with data provided by the public SPIE-AAPM-NCI Prostate MR Classification (PROSTATEx) Challenge. This dataset includes a retrospective study of 204 patients for the training cohort, that correspond to MP-MRi studies with a total of 320 labeled lesions. For the testing cohort, the dataset is compounded by 120 studies that correspond to 210 lesions. For each of these lesions, the dataset provides radiological label findings related to the localization of the lesions and the clinical significance of the lesions supported by the biopsy test.

Each study counts with four sets of MRI sequences: two sets of T2-weighted images, ADC and B-VAL images computed from DWI, and K^{trans} images (computed from dynamic contrast-enhanced (DCE) images)[96]. In clinical routine such modalities complement radi-

[96] Geert Litjens et al. “Computer-aided detection of prostate cancer in MRI”. in: *IEEE transactions on medical imaging* 33.5 (2014), pp. 1083–1092.

ologist analysis and observations, integrating information related to morphology patterns, zonal information, and abnormal activation related to malignant tumors. For instance, the T2WI sequences offer anatomical lesion information. Then, the information available from DWI sequences (ADC and B-value) related to the restriction of water diffusion, may complement lesion discrimination with respect to high cellular density regions. In fact, some state-of-the-art studies[97] have reported a correlation between the quantitative ADC maps and, the Gleason Score (a measure of prostate cancer aggressiveness used mainly in biopsies). For such a reason, the DWI is considered a complement for T2WI to discriminate potential carcinoma lesions in the prostate gland and contribute to the assessment of the disease. in Figure 11, it can be observed a multi-parametric analysis of a malignant lesion.

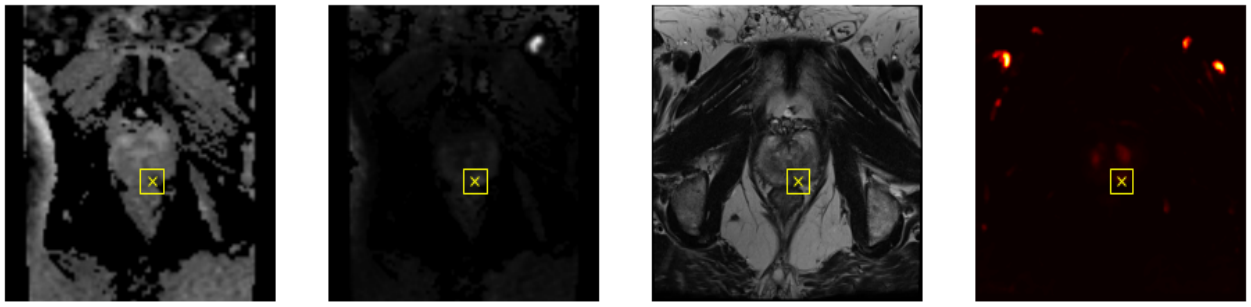


Figure 11. A malignant prostate lesion located in the peripheral zone observed from the multi-parametric perspective of MRI sequences. From left to right: ADC, Maximum B Value, T2WI, and K^{trans} sequences.

These MRI sequences were acquired using two 3T magnetic resonance scanners, the Siemens MAGNETOM Trio and Skyra under the following configurations: T2WI images were acquired using a turbo spin echo sequence configuration with a resolution around

---

[97] Li Zhang et al. "The utility of diffusion MRI with quantitative ADC measurements for differentiating high-grade from low-grade cerebral gliomas: evidence from a meta-analysis". In: *Journal of the neurological sciences* 373 (2017), pp. 9–15; Yu-Chuan Hu et al. "Comparison between ultra-high and conventional mono b-value DWI for preoperative glioma grading". In: *Oncotarget* 8.23 (2017), p. 37884; SD Chen et al. "The correlation between MR diffusion-weighted imaging and pathological grades on glioma". In: *Eur Rev Med Pharmacol Sci* 18.13 (2014), pp. 1904–1909.

0.5 mm in-plane, and a slice thickness of 3.6 mm. The DWI sequences were obtained with a single-shot echo-planar imaging configuration with a resolution of 2 mm in-plane and 3.6 mm slice thickness and with diffusion-encoding gradients in three directions using three b-values intensities (50, 400, and 800). Subsequently, the ADC maps were automatically estimated by the scanners software[98].

5.4.2. Contrastive learning framework setup A crucial point in a weakly supervised scheme is to define input tuples to carry out contrastive learning. In this work, we center a 2D Gaussian ($\mu = 0$) around each radiologist annotation with variance $\sigma = [0.0625, 0.25)$ and over T2WI-MRI sequences. Patches centers in each distribution were taken as malignant samples. The control patches were then taken outside of this Gaussian region but inside of prostate gland segmentation. These samples were randomly collected with the rule that among patches should have a minimum euclidean distance $d = 40$ pixels. Each patch was cropped as volumetric information of $(12 \times 32 \times 32)$.

Regarding the deep architecture, we designed a 3D convolutional neural network f_θ inspired on the Mehrtash's architecture[99]. This deep model characterizes MRI regions as three independent convolutional branches, compounded by nine 3D convolutional layers using the LeakyReLu ($a = 0.3$) activation function. The architecture integrated the zonal information using a one-hot encoded embedding vector, fused at the end of the convolutional branches.

During the training, each volumetric MRI region was represented with a fused embedding feature vector $f_\theta(x_i) \in \mathbf{R}^{128}$. The optimization follows an RMSprop algorithm using the momentum of 0.6 and a learning rate of 1×10^{-6}.

---

98 Litjens et al., "Computer-aided detection of prostate cancer in MRI".

99 Mehrtash et al., "Classification of clinical significance of MRI prostate findings using 3D convolutional neural networks".

Additionally, to measure the textural similarities among MRI regions, the proposed approach underlies on a contrastive learning scheme that was validated using the Triplet Loss (TL), and the NT-Xent loss function $L^{i,j}$. From a validation cohort (sub-set of data extracted from the training cohort) the τ parameter was fixed in $\tau = 0.07$. Also, to increase the amount of MRI regions, we artificially augmented the samples through image transformations such as random rotations, flipping with respect to the horizontal plane, and horizontal and vertical translations.

A main limitation to transfer support technologies is the scarce availability of redundant, and balanced data in clinical centers. For such a reason in this work, we validated the proposed approach with different challenging scenarios that only take a reduced amount of training data. In this case, we sub-sampled the original dataset from $20-100\%$, keeping the same proportion of lesions per class as the original dataset. Additionally, all the reported experiments were obtained from the test set defined and evaluated by the authors of the PROSTATEx challenge. The metric performance selected by the challenge is the Area Under the Receiver Operating Characteristic Curve (ROC-AUC).

6. EVALUATION AND RESULTS

An exhaustive validation of the proposed approach was here carried out with respect to the capability to discriminate prostate cancer lesions from control lesions. The whole experiments report test scores provided by the public PROSTATEx challenge. A first experiment was carried out to evaluate the performance of the proposed contrastive learning scheme, with respect to different strategies to select candidate patches. In such a case, four different strategies were considered to select relevant MRI patches during the weakly supervised stage.

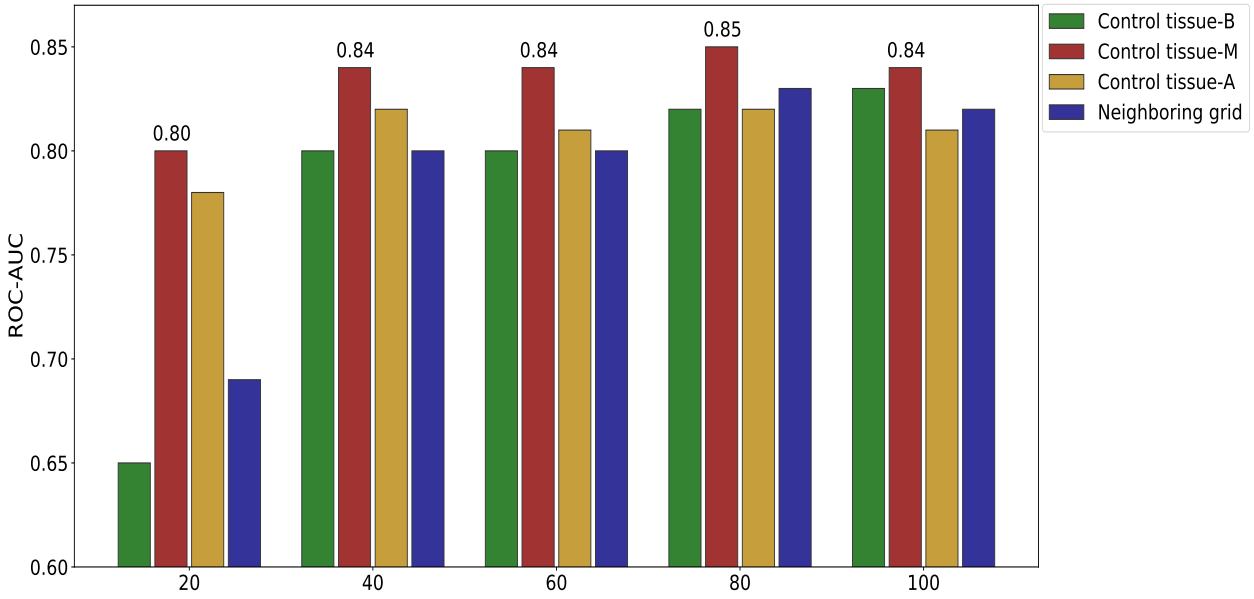


Figure 12. ROC-AUC performance achieved by different selection strategies to extract relevant candidate MRI regions.

The first strategy corresponds to the selection of eight candidate regions from neighboring regions that surround the image-level radiologist annotations (neighboring grid). The second strategy (Control tissue-A) consists of the selection of key control tissue patches $H(x^k)$ estimated from the prostate segmentation tissue $U_\theta(x^k_{T2WI})$, but excluding the lesion affected region maps $G(x^k_{T2WI})$ centered in the radiologist annotations. Additionally,

for strategies three (Control tissue-M) and four (Control tissue-B), we performed a similar strategy but included only the benignant (B) and malignant (M) confirmed lesions respectively. These configurations were run following the NT-Xent loss function and the triplet loss.

Loss function	**Neighboring grid**	**Control tissue**		
		All	**Malign**	**Benign**
NT Xent	0.82	0.81	**0.84**	0.83
Triplet loss	0.81	0.81	**0.82**	0.79

Table 1. Comparison of different candidate selection strategies using the NT-Xent and the Triplet Loss objective functions over all the labeled MRI data.

Table 1 summarizes the achieved results for the different strategies to select candidate patches regarding two contrastive loss functions. The patches selected from malignant lesions as reference achieved the best performance (0.84 from NT Xent, and 0.82 from triple loss). Such selection has a gain of about 2% with respect to the neighboring grid, being interesting to have a reference with respect to some lesions to recover the other control patches. It should be noted, that a careful delineation of lesion regions allows to better characterization and discrimination of patches, which in consequence determines a proper deep representation of lesions. Also, the NT-Xent was superior in whole experiments regarding the triplet loss. This could be attributed to the nature of the NT-Xent loss, which globally estimates the textural variability among an anchor region and N negative regions.

Interestingly, a second experiment evaluated the capability of the weakly contrastive approach to capture complex textural lesion patterns from a reduced set of samples. For doing so, we run several experiments using random subsets taking incremental parts of training data (each 20% of available data). Figure 12 summarizes the achieved results in this experiment, finding consistent results from the Control tissue-M strategy that achieved the best results in all training subsets. It should be noted that the proposed approach achieves a competitive ROC-AUC of 0.84 using only 40% of the total training set, which

suggests a potential use on real scenarios with scarce label sets. Also, the best result was achieved with 80% of total data in training, achieving a remarkable ROC-AUC of 0.85. It should be noted that strategy shows a robust characterization of malignant lesions, with the capability to deal with complex textural observations of the lesions. This fact may be associated with the weakly strategy that approaches other regions of the images to adjust deep representations.

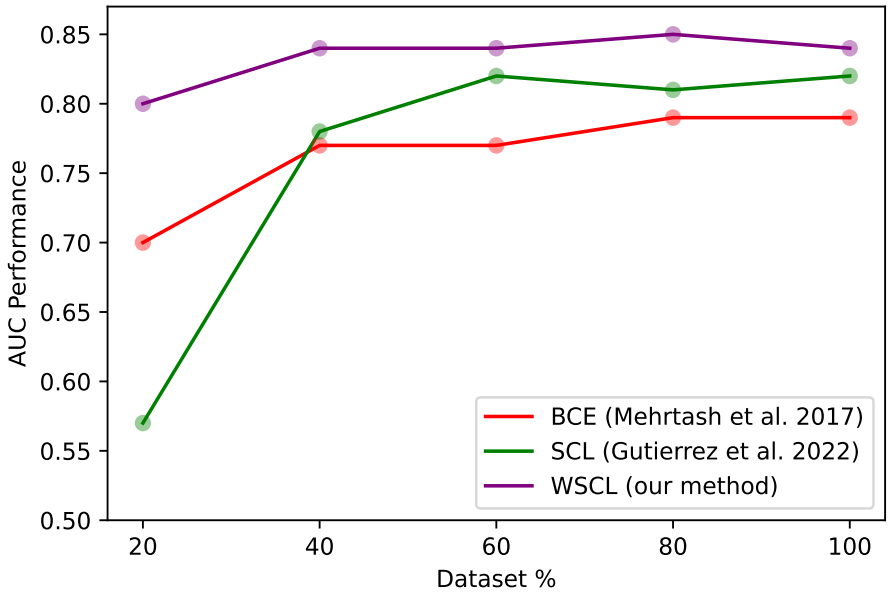


Figure 13. Results obtained by our proposed WSCL of prostate regions in MRI, and baseline supervised learning schemes under different data configurations. In red, we have the work proposed by Mehrtash using a BCE loss[100], in green, we have our multimodal supervised contrastive learning approach[101], and finally in purple, we have our WSCL scheme.

As a baseline, we implemented a standard architecture validated in this challenge[102]. Hence, we trained and adjusted this representation following a classical binary cross-

[102] Mehrtash et al., "Classification of clinical significance of MRI prostate findings using 3D convolutional neural networks".

entropy (BCE) loss function[103], following a supervised contrastive learning (SCL) scheme[104] and also with the here introduced weakly supervised strategy (WSCL). To evidence the performance of the learning approaches with scarce training scenarios, we run several experiments with incremental subsets of labeled MRI lesions to measure the capability of these approaches under clinical scenarios with the scarcity of labeled MRI data. As observed in figure 13, the proposed WSCL scheme (Control tissue-M + XT-Xent) obtained the labeled lesions, and our WSCL achieved the best ROC-AUC results for all of the available data configurations. Remarkably, using only 20% of the labeled lesions achieved a ROC-AUC of 0.80, obtaining an improvement of 10 and 24% with respect to the reported baseline strategies. It should be also noted, that contrastive schemes were the best alternative to train schemes in scenarios with scarce data, while classical cross-entropy rules remain limited to learn lesion variability.

A qualitative analysis was here carried out by recovering salience maps that stand out regions, at each MRI parameter, that major contributes to the final prediction. To recover such attention maps, we run the GradCAM strategy over the selected backbone after the trained representation. This strategy back propagated the output prediction into convolutional branches, allowing to evidence localized regions with major association with the estimations. Figure 14 shows the retrieved maps for BVAL and ADC maps at different labeled data configurations. As expected, there is a general major activation of these maps in malignant regions. Interestingly enough, this is coherent with respect to previous studies that support a negative correlation between the Gleason Grade of Gliomas and the DWI sequence. On the other hand, as observed in figure 14 the T2WI sequence complements

---

[103] Mehrtash et al., “Classification of clinical significance of MRI prostate findings using 3D convolutional neural networks”.

[104] Gutiérrez, Arevalo, and Martínez, “Multimodal Contrastive Supervised Learning to Classify Clinical Significance MRI Regions on Prostate Cancer”.

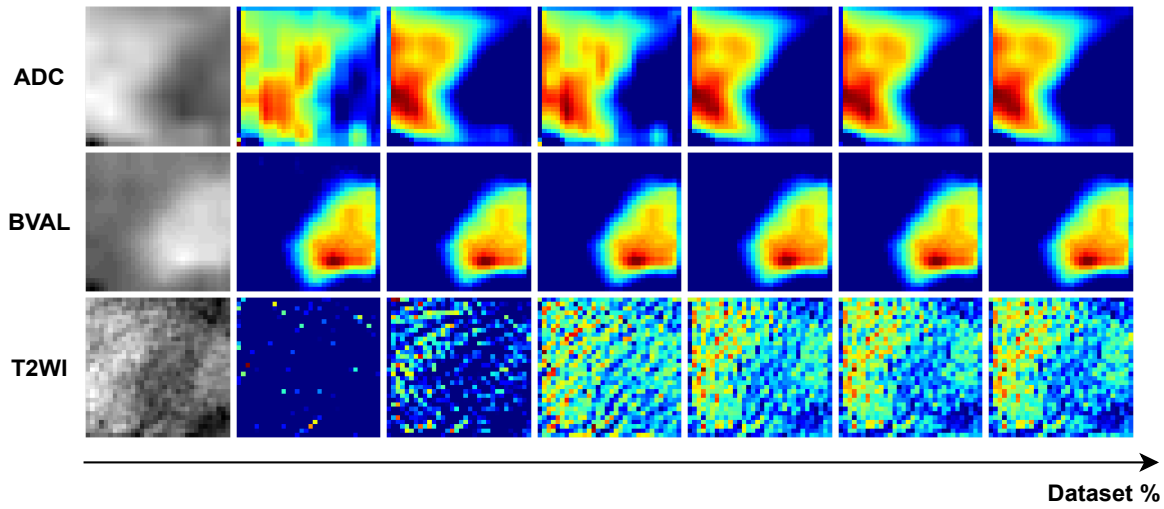


Figure 14. GradCAM attention maps obtained by the deep visual representation over a malignant lesion; from top to bottom we have ADC, BVAL, and T2WI sequences. From left to right we have the progression over the attention maps for the weakly pre-trained representation fine-tuned over 0, 20, 40, 60, 80, and 100% of the labeled dataset respectively.

the characterization depicting textural properties related to the anatomy of the prostatic tissue of study. Additionally, in the second column of the figure, it can be observed how most of the attention maps are already learned from the pre-trained representation that discriminates lesion and control regions. Then from left to right, it is possible to evidence how for each of the MRI sequences, the attention maps are fine-tuned by progressively including more data to adjust and learn the resultant deep representation.

7. DISCUSSION

This work presented a BP-MRI Weakly Supervised Contrastive Learning (WSCL) strategy that discriminates malignant lesions related to prostate cancer disease from suspicious image-level annotations. The proposed strategy was adjusted, measuring the mutual information among MRI samples of prostate tissue, and dealing with the scarcity of labels, as typically reported in clinical scenarios. The proposed approach outperforms standard learning approaches, achieving a ROC-AUC of 85%, and using only 80% of total data. Remarkably, the proposed approach uses only 20% of available training data and reports a ROC-AUC of 80%.

The proposed approach significantly increases the amount of training data by following the weakly supervised strategy, allowing a first coarse classification between benign and malignant tissue. As supported in Figure 14, this first training stage is sufficient to recover abnormal tissue patterns, from BP-MRI studies. Then, in the second training stage, the fine-tuning allows more sensible discrimination among abnormal tissues related to cancer, regarding other prostate affectations. As expected, the maps encoded from DWI such as BVAL and the ADC seem to be the most important MRI sequences to characterize malignant lesions, this is an interesting fact that has also been reported in the literature since DWI maps have shown a correlation with the Gleason grade of gliomas[105] [106] [107]. Consequently, we decided to principally adjust the representation to encode patterns from

---

[105] Zhang et al., "The utility of diffusion MRI with quantitative ADC measurements for differentiating high-grade from low-grade cerebral gliomas: evidence from a meta-analysis".

[106] Hu et al., "Comparison between ultra-high and conventional mono b-value DWI for preoperative glioma grading".

[107] Chen et al., "The correlation between MR diffusion-weighted imaging and pathological grades on glioma".

a bi-parametric (BP-MRI) perspective (integrating only T2WI and DWI), according to recently suggested in PI-RADS v2 protocol, avoiding the contrast agent dependency[108] [109]. In the literature, exist different solutions to support malignant lesion classification from MP-MRI studies.

For instance, Mehrtash's work proposed a 3D convolutional backbone that learns independent paths for each parameter, which is further integrated into an embedding representation. Also, early fusion strategies from inception modules have been implemented to classify prostate cancer lesions. These approaches nonetheless lie on the classical cross-entropy rule with strictly supervised schemes, that may collapse representation, overfitting specific patterns and limiting deep representation to support classification. Contrary, the proposed approach using a restricted set of parameters (bi-parametric representation) overcame the ROC-AUC diagnosis of these strategies by 10 and 20% respectively using only the 20% of the labeled MRI data (see in Figure 13). Since these three strategies share a similar pipeline in terms of deep architecture, image pre-processing, and cropping of MRI regions, we hypothesize that the obtained results are due to our proposed WSCL scheme that not only takes into account the regions labeled by expert radiologists but also the pseudo-labeled control MRI regions that were estimated from the prostate gland.

To support class imbalance, Chen et al[110] proposed a Transfer Learning scheme from the open ImageNet dataset to adjust the deep representation into the clinical domain. As a result, the author obtained a ROC-AUC of 0.82 using 100% of the annotated data.

---

[108] Murphy, "The expanding role of MRI in prostate cancer".

[109] Barentsz et al., "ESUR prostate MR guidelines 2012".

[110] Chen et al., "A transfer learning approach for classification of clinical significant prostate cancers from mpMRI scans".

Additionally, other studies in the state-of-the-art such as Liu et al[111], and Aldoj et al[112] were also proposed to characterize malignant regions in the prostate gland using MRI sequences, achieving a ROC-AUC of 0.84 and 0.89 respectively. Nevertheless, these authors included different data architectures, image pre-processing pipelines, and cropping region schemes, which difficult a proper comparison against our proposed strategy. Moreover, these authors didn't report more data configurations to evaluate their performance in clinical scenarios with a scarcity of data.

The proposed strategy achieves competitive results, reaching remarkable results to face challenging scenarios with scarce labeled data for training. We hypothesize that weakly supervised learning strategies could empower deep learning applications in clinical scenarios by adjusting deep architectures directly with clinical data, and therefore facilitate the technological transfer of these strategies to the clinical routine. Despite important advances in such representations, it is expected to evolve the proposed model to totally avoid dependency on labeled annotations, as well as, to conduct experiments over larger test datasets to validate the generalization capabilities of the proposed approach.

---

[111] Liu et al., "Prostate cancer diagnosis using deep learning with 3D multiparametric MRI".

[112] Aldoj et al., "Semi-automatic classification of prostate cancer on multi-parametric MR imaging using a multi-channel 3D convolutional neural network".

8. CONCLUSIONS AND FUTURE WORK

This work introduced a weakly supervised contrastive learning strategy to characterize malignant MRI regions related to prostate cancer disease. The proposed strategy encodes multimodal morphological and cellular density patterns available in T2WI and DWI (specifically from ADC and B-VAL maps) to represent control and lesion regions in a contrastive embedded space. A fine-tuning stage allowed cancer lesion discrimination from a linear hyperplane. The achieved results suggest that contrasting prostate lesions with control regions, under a weakly supervised learning scheme improved the characterization and discrimination of malignant lesions, especially in clinical scenarios with a limited amount of labeled lesions. Moreover, our experimental setup shows that the deep representation obtained a decent ROC-AUC performance of 0.8 in the diagnosis using only the 20% of annotated lesions. As a result, the deep representation better differentiates between healthy and lesion regions in a projected contrastive embedded space. This WSCL strategy could be potentially used in clinical scenarios with scarce of annotated MRI data to empower deep learning applications in the clinical stages and support the diagnosis of the disease. Future works include the evaluation over extensive datasets, from different scanners, allowing the establishment of a generalization level of the weakly supervised strategy. Also, there will be carried out different efforts to avoid human labeled annotations during training, through the development of self-supervised mechanisms.

BIBLIOGRAPHY

Ahmed, Hashim Uddin et al. *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.* 2016.

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

As, Nicholas J van et al. “A study of diffusion-weighted magnetic resonance imaging in men with untreated localised prostate cancer on active surveillance”. In: *European urology* 56.6 (2009), pp. 981–988.

Azimi, Seyed Majid et al. “Towards multi-class object detection in unconstrained remote sensing imagery”. In: *Asian Conference on Computer Vision*. Springer. 2018, pp. 150–165.

Barentsz, Jelle O et al. “ESUR prostate MR guidelines 2012”. In: *European radiology* 22.4 (2012), pp. 746–757.

Barry, Michael J. “Prostate-specific–antigen testing for early diagnosis of prostate cancer”. In: *New England Journal of Medicine* 344.18 (2001), pp. 1373–1377.

Berman, Rose M et al. “DCE MRI of prostate cancer”. In: *Abdominal Radiology* 41.5 (2016), pp. 844–853.

Blake, Michael A and Mannudeep K Kalra. *Imaging in oncology*. Vol. 143. Springer Science & Business Media, 2008.

Bleker, Jeroen et al. “Multiparametric MRI and auto-fixed volume of interest-based radiomics signature for clinically significant peripheral zone prostate cancer”. In: *European radiology* 30.3 (2020), pp. 1313–1324.

Chan, Ian et al. “Detection of prostate cancer by integration of line-scan diffusion, T2-mapping and T2-weighted magnetic resonance imaging; a multichannel statistical classifier”. In: *Medical physics* 30.9 (2003), pp. 2390–2398.

Chen, Quan et al. “A transfer learning approach for classification of clinical significant prostate cancers from mpMRI scans”. In: *Medical Imaging 2017: Computer-Aided Diagnosis*. Vol. 10134. International Society for Optics and Photonics. 2017, 101344F.

Chen, SD et al. “The correlation between MR diffusion-weighted imaging and pathological grades on glioma”. In: *Eur Rev Med Pharmacol Sci* 18.13 (2014), pp. 1904–1909.

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

Chen, Ting et al. “Big self-supervised models are strong semi-supervised learners”. In: *arXiv preprint arXiv:2006.10029* (2020).

Ciga, Ozan, Anne L Martel, and Tony Xu. “Self supervised contrastive learning for digital histopathology”. In: *arXiv preprint arXiv:2011.13971* (2020).

Clements, R et al. “Side effects and patient acceptability of transrectal biopsy of the prostate”. In: *Clinical radiology* 47.2 (1993), pp. 125–126.

Crawford, E David et al. “Serum prostate-specific antigen and digital rectal examination for early detection of prostate cancer in a national community-based program”. In: *Urology* 47.6 (1996), pp. 863–869.

Cuenod, CA and D. Balvay. “Perfusion and vascular permeability: basic concepts and measurement in DCE-CT and DCE-MRI”. In: *Diagnostic and interventional imaging* 94.12 (2013), pp. 1187–1204.

Essig, Marco et al. “Perfusion MRI: the five most frequently asked technical questions”. In: *American Journal of Roentgenology* 200.1 (2013), pp. 24–34.

Ferlay, J et al. “Estimating the global cancer incidence and mortality in 2018: GLOBOCAN sources and methods”. In: *International Journal of Cancer* 144 (2018).

Gaur, Sonia et al. “Can computer-aided diagnosis assist in the identification of prostate cancer on prostate MRI? a multi-center, multi-reader investigation”. In: *Oncotarget* 9.73 (2018), p. 33804.

Giannini, Valentina et al. “A fully automatic computer aided diagnosis system for peripheral zone prostate cancer detection using multi-parametric magnetic resonance imaging”. In: *Computerized Medical Imaging and Graphics* 46 (2015), pp. 219–226.

Girouin, Nicolas et al. “Prostate dynamic contrast-enhanced MRI with simple visual diagnostic criteria: is it reasonable?” In: *European radiology* 17.6 (2007), pp. 1498–1509.

Greene, D, A Ali, N Kinsella, et al. *Transrectal Ultrasound and Prostatic Biopsy: Guidelines & Recommendations for Training*. 2015.

Greer, Matthew D et al. “Interreader variability of prostate imaging reporting and data system version 2 in detecting and assessing prostate cancer lesions at prostate MRI”. In: *AJR. American journal of roentgenology* (2019), p. 1.

Gudmundsson, Julius et al. “Genome-wide associations for benign prostatic hyperplasia reveal a genetic correlation with serum levels of PSA”. In: *Nature communications* 9.1 (2018), pp. 1–8.

Gupta, Rajan T et al. “Multiparametric prostate MRI: focus on T2-weighted imaging and role in staging of prostate cancer”. In: *Abdominal Radiology* 41.5 (2016), pp. 831–843.

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

Gutiérrez, Yesid, John Arevalo, and Fabio Martıénez. “A Ktrans deep characterization to measure clinical significance regions on prostate cancer”. In: *15th International Symposium on Medical Information Processing and Analysis*. Vol. 11330. SPIE. 2020, pp. 80–88.

Gutiérrez, Yesid, John Arevalo, and Fabio Martıénez. “An inception-based deep multiparametric net to classify clinical significance MRI regions of prostate cancer”. In: *Physics in Medicine & Biology* 67.22 (2022), p. 225004.

Gutiérrez, Yesid, Gustavo Garzón, and Fabio Martıénez. “Towards clinical significance prediction using k trans evidences in prostate cancer”. In: *2019 XXII Symposium on Image, Signal Processing and Artificial Vision (STSIVA)*. IEEE. 2019, pp. 1–5.

Hajian-Tilaki, Karimollah. “Receiver operating characteristic (ROC) curve analysis for medical diagnostic test evaluation”. In: *Caspian journal of internal medicine* 4.2 (2013), p. 627.

Hambrock, Thomas et al. “Relationship between apparent diffusion coefficients at 3.0-T MR imaging and Gleason grade in peripheral zone prostate cancer”. In: *Radiology* 259.2 (2011), pp. 453–461.

Hamoen, Esther HJ et al. “Use of the prostate imaging reporting and data system (PI-RADS) for prostate cancer detection with multiparametric magnetic resonance imaging: a diagnostic meta-analysis”. In: *European urology* 67.6 (2015), pp. 1112–1121.

Hara, Noboru et al. “Dynamic contrast-enhanced magnetic resonance imaging (DCE-MRI) is a useful modality for the precise detection and staging of early prostate cancer”. In: *The Prostate* 62.2 (2005), pp. 140–147.

He, Kaiming et al. “Deep residual learning for image recognition”. In: *Proceedings of the IEEE conference on computer vision and pattern recognition*. 2016, pp. 770–778.

He, Kaiming et al. “Momentum contrast for unsupervised visual representation learning”. In: *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. 2020, pp. 9729–9738.

Holmström, Benny et al. “Prostate specific antigen for early detection of prostate cancer: longitudinal study”. In: *Bmj* 339 (2009).

Hötker, Andreas M et al. “Assessment of prostate cancer aggressiveness by use of the combination of quantitative DWI and dynamic contrast-enhanced MRI”. In: *American Journal of Roentgenology* 206.4 (2016), pp. 756–763.

Hu, Yu-Chuan et al. “Comparison between ultra-high and conventional mono b-value DWI for preoperative glioma grading”. In: *Oncotarget* 8.23 (2017), p. 37884.

Kasivisvanathan, Veeru et al. “MRI-targeted or standard biopsy for prostate-cancer diagnosis”. In: *New England Journal of Medicine* 378.19 (2018), pp. 1767–1777.

Khosla, Prannay et al. “Supervised contrastive learning”. In: *arXiv preprint arXiv:2004.11362* (2020).

Labrie, Fernand et al. “Screening decreases prostate cancer death: first analysis of the 1988 Quebec prospective randomized controlled trial”. In: *The Prostate* 38.2 (1999), pp. 83–91.

Langer, Deanna L et al. “Prostate cancer detection with multi-parametric MRI: Logistic regression analysis of quantitative T2, diffusion-weighted imaging, and dynamic contrast-enhanced MRI”. In: *Journal of Magnetic Resonance Imaging: An Official Journal of the International Society for Magnetic Resonance in Medicine* 30.2 (2009), pp. 327–334.

Le, Minh Hung et al. “Automated diagnosis of prostate cancer in multi-parametric MRI based on multimodal convolutional neural networks”. In: *Physics in Medicine & Biology* 62.16 (2017), p. 6497.

LeCun, Yann et al. “Gradient-based learning applied to document recognition”. In: *Proceedings of the IEEE* 86.11 (1998), pp. 2278–2324.

Li, Jie et al. “Angiogenesis in wound repair: angiogenic growth factors and the extracellular matrix”. In: *Microscopy research and technique* 60.1 (2003), pp. 107–114.

Litjens, Geert et al. “Computer-aided detection of prostate cancer in MRI”. In: *IEEE transactions on medical imaging* 33.5 (2014), pp. 1083–1092.

Liu, Saifeng et al. “Prostate cancer diagnosis using deep learning with 3D multiparametric MRI”. In: *Medical imaging 2017: computer-aided diagnosis*. Vol. 10134. International Society for Optics and Photonics. 2017, p. 1013428.

Loeb, Stacy et al. “Overdiagnosis and overtreatment of prostate cancer”. In: *European urology* 65.6 (2014), pp. 1046–1055.

Manley, Brandon J et al. “Prostate MRI: a national survey of Urologist’s attitudes and perceptions”. In: *International braz j urol* 42.3 (2016), pp. 464–471.

McInnes, Leland, John Healy, and James Melville. “Umap: Uniform manifold approximation and projection for dimension reduction”. In: *arXiv preprint arXiv:1802.03426* (2018).

Mehrtash, Alireza et al. “Classification of clinical significance of MRI prostate findings using 3D convolutional neural networks”. In: *Medical Imaging 2017: Computer-Aided Diagnosis*. Vol. 10134. International Society for Optics and Photonics. 2017, 101342A.

Mettlin, Curtis et al. “The American cancer society national prostate cancer detection project. Findings on the detection of early prostate cancer in 2425 men”. In: *Cancer* 67.12 (1991), pp. 2949–2958.

Meyer, Anneke et al. “Towards patient-individual PI-Rads v2 sector map: CNN for automatic segmentation of prostatic zones from T2-weighted MRI”. In: *2019 IEEE 16th International Symposium on Biomedical Imaging (ISBI 2019)*. IEEE. 2019, pp. 696–700.

Miao, Huadong, Hiroshi Fukatsu, and Takeo Ishigaki. “Prostate cancer detection with 3-T MRI: comparison of diffusion-weighted and T2-weighted imaging”. In: *European journal of radiology* 61.2 (2007), pp. 297–302.

Murphy et. al, G. “The expanding role of MRI in prostate cancer”. In: *American Journal of Roentgenology* 201.6 (2013), pp. 1229–1238.

Neal Jr, Durwood E et al. “Prostate specific antigen and prostatitis I. Effect of prostatitis on serum PSA in the human and nonhuman primate”. In: *The Prostate* 20.2 (1992), pp. 105–111.

O'Connor, James PB et al. "DCE-MRI biomarkers in the clinical evaluation of antiangiogenic and vascular disrupting agents". In: *British journal of cancer* 96.2 (2007), pp. 189–195.

Oord, Aaron van den, Yazhe Li, and Oriol Vinyals. "Representation learning with contrastive predictive coding". In: *arXiv preprint arXiv:1807.03748* (2018).

Oto, Aytekin et al. "Diffusion-weighted and dynamic contrast-enhanced MRI of prostate cancer: correlation of quantitative MR parameters with Gleason score and tumor angiogenesis". In: *American Journal of Roentgenology* 197.6 (2011), pp. 1382–1390.

Penn, John. *Retinal and choroidal angiogenesis*. Springer Science & Business Media, 2008.

Polascik, Thomas J, Joseph E Oesterling, and Alan W Partin. "Prostate specific antigen: a decade of discovery-what we have learned and where we are going". In: *The Journal of urology* 162.2 (1999), pp. 293–306.

Porter, Kristin K et al. "Financial implications of biparametric prostate MRI". In: *Prostate cancer and prostatic diseases* 23.1 (2020), pp. 88–93.

Quinlan, Mark R, Damien Bolton, and Rowan G Casey. "The management of rectal bleeding following transrectal prostate biopsy: A review of the current literature". In: *Canadian Urological Association Journal* 12.3 (2018), E146.

Rooij, Maarten de et al. "Accuracy of multiparametric MRI for prostate cancer detection: a meta-analysis". In: *American Journal of Roentgenology* 202.2 (2014), pp. 343–351.

Rossi, Alberto, Monica Bianchini, and Franco Scarselli. "Robust prostate cancer classification with siamese neural networks". In: *International Symposium on Visual Computing*. Springer. 2020, pp. 180–189.

Saar, Matthias et al. "Current role of multiparametric MRI and MRI targeted biopsies for prostate cancer diagnosis in Germany: a nationwide survey". In: *Urologia Internationalis* 104.9-10 (2020), pp. 731–740.

Schroff, Florian et al. “Facenet: A unified embedding for face recognition and clustering”. In: *Proceedings of the IEEE conference on computer vision and pattern recognition*. 2015, pp. 815–823.

Serefoglu, Ege Can et al. “How reliable is 12-core prostate biopsy procedure in the detection of prostate cancer?” In: *Canadian Urological Association Journal* 7.5-6 (2013), E293.

Shahait, Mohammed et al. “Incidence of sepsis following transrectal ultrasound guided prostate biopsy at a tertiary-care medical center in Lebanon”. In: *International braz j urol* 42.1 (2016), pp. 60–68.

Siegel et al., R. L. “Cancer statistics, 2020”. In: *CA: a cancer journal for clinicians* 70.1 (2013), pp. 7–30.

Smith, Deborah S and William J Catalona. “Interexaminer variability of digital rectal examination in detecting prostate cancer”. In: *Urology* 45.1 (1995), pp. 70–74.

Sohn, Kihyuk. “Improved deep metric learning with multi-class n-pair loss objective”. In: *Advances in neural information processing systems*. 2016, pp. 1857–1865.

Sukhbaatar, Sainbayar et al. “Training convolutional networks with noisy labels”. In: *arXiv preprint arXiv:1406.2080* (2014).

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

Sunoqrot, Mohammed RS et al. “Automated reference tissue normalization of T2-weighted MR images of the prostate using object recognition”. In: *Magnetic Resonance Materials in Physics, Biology and Medicine* (2020), pp. 1–13.

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.

Tsehay, Yohannes K et al. “Convolutional neural network based deep-learning architecture for prostate cancer detection on multiparametric magnetic resonance images”. In: *Medical imaging 2017: Computer-aided diagnosis*. Vol. 10134. SPIE. 2017, pp. 20–30.

Vargas, HA et al. “Updated prostate imaging reporting and data system (PIRADS v2) recommendations for the detection of clinically significant prostate cancer using multiparametric MRI: critical evaluation using whole-mount pathology as standard of reference”. In: *European radiology* 26.6 (2016), pp. 1606–1612.

Verma, Sadhna et al. “Overview of dynamic contrast-enhanced MRI in prostate cancer diagnosis and management”. In: *American Journal of Roentgenology* 198.6 (2012), pp. 1277–1288.

Wang, Zhiwei et al. “Automated detection of clinically significant prostate cancer in mp-MRI images based on an end-to-end deep neural network”. In: *IEEE transactions on medical imaging* 37.5 (2018), pp. 1127–1139.

Woodfield, Courtney A et al. “Diffusion-weighted MRI of peripheral zone prostate cancer: comparison of tumor apparent diffusion coefficient with Gleason score and percentage of tumor on core biopsy”. In: *American Journal of Roentgenology* 194.4 (2010), W316–W322.

Xuan et. al, H. “Improved embeddings with easy positive triplet mining”. In: *Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision*. 2020, pp. 2474–2482.

Yang, Xin et al. “Co-trained convolutional neural networks for automated detection of prostate cancer in multi-parametric MRI”. In: *Medical image analysis* 42 (2017), pp. 212–227.

Zhang, Angela, Thomas Fear, and Hammood Ahmed. “Digital rectal examination in prostate cancer screening”. In: *University of Western Ontario Medical Journal* 82.1 (2013), pp. 10–11.

Zhang, Li et al. “The utility of diffusion MRI with quantitative ADC measurements for differentiating high-grade from low-grade cerebral gliomas: evidence from a meta-analysis”. In: *Journal of the neurological sciences* 373 (2017), pp. 9–15.

Zhang, Na et al. “Correlation of volume transfer coefficient Ktrans with histopathologic grades of gliomas”. In: *Journal of Magnetic Resonance Imaging* 36.2 (2012), pp. 355–363.

Zhang, Zhilu and Mert R Sabuncu. “Generalized cross entropy loss for training deep neural networks with noisy labels”. In: *arXiv preprint arXiv:1805.07836* (2018).

*

APPENDIX

appendix A. Academic Products

Journals

- 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.
- Gutiérrez, Y., Arevalo, J., & Martínez, F. (2023). A contrastive weakly supervised learning to characterize malignant prostate lesions in BP-MRI.
  Submitted for review at the computer methods and programs in Biomedicine Journal.

Conference papers

- 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.
- Gutiérrez, Y., Olmos J., Guayacán L., & Martínez F. A multimodal geometric deep representation to support bi-parametric prostate cancer lesion classification.
  Conference paper accepted at 20th IEEE International Symposium on Biomedical Imaging (ISBI 2023)