Computer-Aided Diagnosis for Prostate Cancer using mp-MRI

PhD Defence
28th November 2016

Guillaume Lemaître

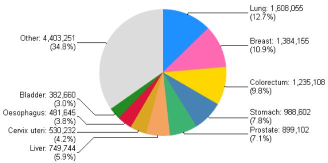
Universitat de Girona - ViCOROB
Université de Bourgogne Franche-Comté - LE2I

Supervised by:

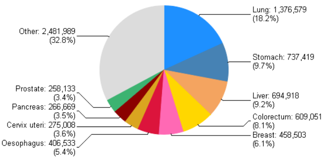
Robert Martí - Fabrice Mériaudeau
Jordi Freixenet - Paul M. Walker

Motivations

Statistics



(a) # of cancer cases



(b) # of cancer deaths

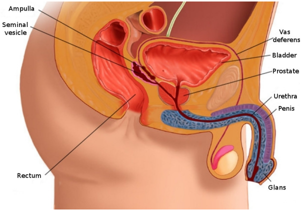
Implications[1]

- 2nd most frequently diagnosed men cancer
- Accounting for 7.1% of overall cancers diagnosed
- Accounting for 3.4% of overall cancers death

[1] J. Ferlay et al. "Estimates of worldwide burden of cancer in 2008: GLOBOCAN 2008". In: *Int. J. Cancer* 127.12 (Dec. 2010), pp. 2893–2917.

The prostate organ

Anatomy



Localization of the prostate organ, image source[2]

Characteristics

- Height: 3 cm
- Depth: 2.5 cm
- Weight: 7 g to 16 g

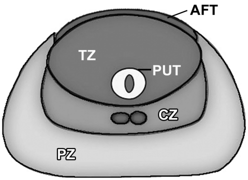
[2] Geckomedia. *Natom Anatomy*. French. June 2011. URL: http://www.natomshop.com/.

UdG

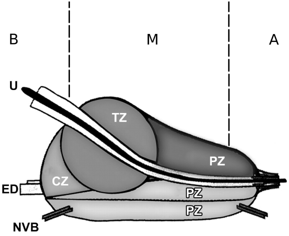
The prostate organ

UBFC

Anatomy



(a) Transverse plane

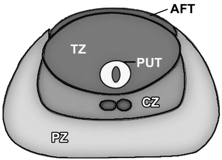


(b) Sagittal plane

Prostate zones - AFT: anterior fibromuscular tissue, CZ: central zone, ED: ejaculatory duct, NVB: neurovascular bundle, PUT: periurethral tissue, PZ: peripheral zone, U: urethra, TZ: transitional zone, B: base, M: median, A: apex; image source[3]

[3] Y. J. Choi et al. "Functional MR imaging of prostate cancer". In: *Radiographics* 27 (2007), pp. 63–75.

Prostate carcinoma (CaP)



CaP development

- Slow-growing $\rightarrow$ 85 %
- Fast-growing $\rightarrow$ 15 %
- CaPs in CG are more aggressive

Zonal predisposition

- PZ $\rightarrow$ 70 % to 80 %
- TZ $\rightarrow$ 10 % to 20 %
- CG $\rightarrow$ 5 %

What clinicians need?

- Detect CaP
- Distinguish slow- from fast-growing CaP
- Active surveillance *vs.* prostatectomy/other treatments

UdG

Screening

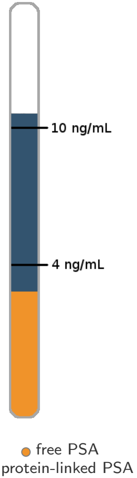
UBFC

Prostate-specific antigen

- $> 10\,\text{ng}\,\text{mL}^{-1} \rightarrow$ biopsy
- From $4\,\text{ng}\,\text{mL}^{-1}$ to $10\,\text{ng}\,\text{mL}^{-1}$
  $\rightarrow \frac{\bullet}{\bullet + \bullet} > 15\% \rightarrow$ biopsy

"Blind" transrectal ultrasound biopsy

- Take samples from different locations
- Grade using Gleason score



UdG

Screening

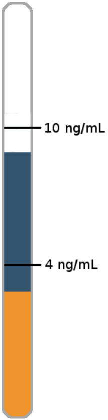
UBFC

Prostate-specific antigen

- $> 10\,\mathrm{ng\,mL^{-1}} \rightarrow$ biopsy
- From $4\,\mathrm{ng\,mL^{-1}}$ to $10\,\mathrm{ng\,mL^{-1}}$
  $\rightarrow \frac{\color{orange}\bullet}{\color{orange}\bullet + \color{blue}\bullet} > 15\% \rightarrow$ biopsy

"Blind" transrectal ultrasound biopsy

- Take samples from different locations
- Grade using Gleason score



free PSA
protein-linked PSA

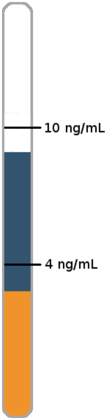
Screening

Prostate-specific antigen

- $> 10\,\text{ng}\,\text{mL}^{-1} \rightarrow$ biopsy
- From $4\,\text{ng}\,\text{mL}^{-1}$ to $10\,\text{ng}\,\text{mL}^{-1}$ $\rightarrow \frac{\bullet}{\bullet + \bullet} > 15\% \rightarrow$ biopsy

"Blind" transrectal ultrasound biopsy

- Take samples from different locations
- Grade using Gleason score



● free PSA
● protein-linked PSA

Screening

Prostate-specific antigen

- $> 10\,\mathrm{ng\,mL^{-1}} \rightarrow$ biopsy
- From $4\,\mathrm{ng\,mL^{-1}}$ to $10\,\mathrm{ng\,mL^{-1}}$
  $\rightarrow \frac{\bullet}{\bullet + \bullet} > 15\% \rightarrow$ biopsy

"Blind" transrectal ultrasound biopsy

- Take samples from different locations
- Grade using Gleason score

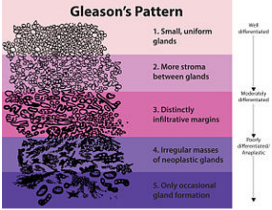


Image source: https://goo.gl/fEVQXQ

UdG

Screening

UBFC

Prostate-specific antigen

- $> 10\,\mathrm{ng\,mL^{-1}} \rightarrow$ biopsy
- From $4\,\mathrm{ng\,mL^{-1}}$ to $10\,\mathrm{ng\,mL^{-1}}$
  $\rightarrow \frac{\color{orange}\bullet}{\color{orange}\bullet + \color{blue}\bullet} > 15\% \rightarrow$ biopsy

"Blind" transrectal ultrasound biopsy

- Take samples from different locations
- Grade using Gleason score

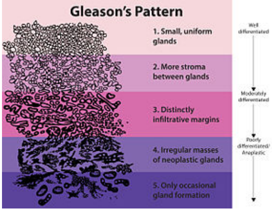


Image source: `https://goo.gl/fEVQXQ`

UdG

Screening

UBFC

Pros

✓ Reduce CaP-related mortality from 21 % to 44 %[4]

Cons

✗ Up to 30 % of over-diagnosis[5]

✗ Up to 35 % of undiagnosed CaP[6]

✗ Biopsies are invasive

[4] Fritz H. Schrder et al. "Prostate-cancer mortality at 11 years of follow-up". In: *New England Journal of Medicine* 366.11 (2012), pp. 981–990.

[5] G. P. Haas et al. "Needle biopsies on autopsy prostates: sensitivity of cancer detection based on true prevalence". In: *J. Natl. Cancer Inst.* 99.19 (Oct. 2007), pp. 1484–1489.

[6] A. V. Taira et al. "Performance of transperineal template-guided mapping biopsy in detecting prostate cancer in the initial and repeat biopsy setting". In: *Prostate Cancer Prostatic Dis.* 13.1 (Mar. 2010), pp. 71–77.

CAD and mp-MRI

Current trendy techniques: mp-MRI

- ✓ Less invasive technique

Human diagnosis using mp-MRI

- ✗ Need further investigation of the mp-MRI modalities
- ✗ Low repeatability
  - Observer limitations
  - Complexity of clinical cases

Emergence of CAD

- CADe $\rightarrow$ detection of potential lesions
- CADx $\rightarrow$ diagnosis regarding those lesions

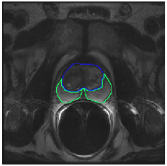
Research objectives

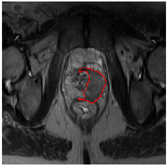
Propose a mp-MRI CAD for CaP

- Study and investigate the state-of-the-art on MRI CAD for CaP
- Identify the scientific barriers
- Design a mp-MRI CAD addressing these issues
- Investigate and analyze the proposed CAD

MRI modalities

T_2W-MRI


(a) Healthy


(b) CaP PZ


(c) CaP CG

Healthy

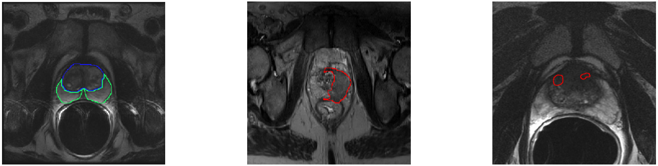
- Intermediate to high-signal intensity (SI) in PZ
- Low-SI in CG

CaP

- Low-SI
- Round and ill-defined mass in PZ
- Homogeneous with ill-defined edges in CG

MRI modalities

T_2W-MRI



(a) Healthy (b) CaP PZ (c) CaP CG

Pros

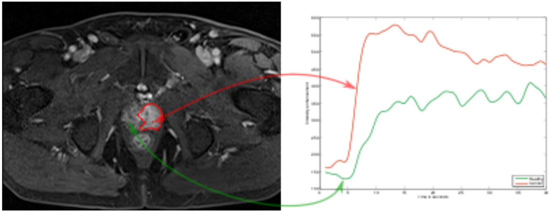
- Highest spatial resolution
- Anatomy nicely depicted

Cons

- Low sensitivity in CG
- Lower specificity due to outliers

MRI modalities

DCE-MRI



Green: healthy - Red: CaP

Healthy

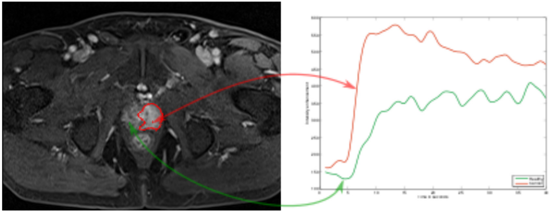
- Slower wash-in, wash-out, time-to-peak enhancement
- Lower integral under the curve, max SI

CaP

- Faster wash-in, wash-out, time-to-peak enhancement
- Higher integral under the curve, max SI

MRI modalities

DCE-MRI



Green: healthy - Red: CaP

Pros

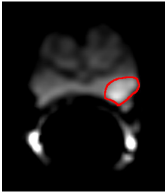
- Information about vascularity

Cons

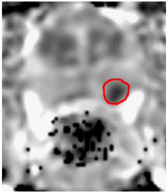
- Spatial mis-registration
- Lower spatial resolution than T_2W-MRI
- Difficult detection in CG

MRI modalities

DW-MRI - ADC



(a) DW MRI



(b) ADC

Healthy

- DW-MRI: lower SI
- ADC: higher-SI

CaP

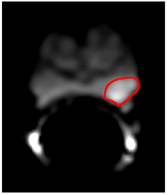
- DW-MRI: higher SI
- ADC: lower-SI

UdG

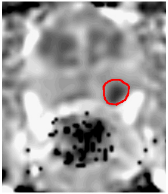
MRI modalities

UBFC

DW-MRI - ADC



(c) DW MRI



(d) ADC

Pros

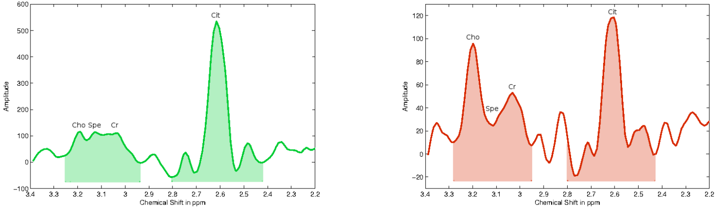
- Information about tissue structure
- ADC correlated with Gleason score

Cons

- Poor spatial resolution
- Variability of the ADC coefficient

MRI modalities

MRSI



(a) Healthy

(b) CaP

Healthy

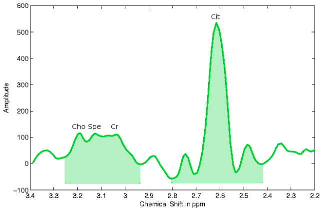
- High citrate concentration
- Moderate choline and spermine concentrations

CaP

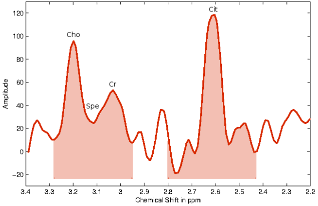
- Decrease of citrate and spermine concentrations
- Increase of choline concentration

MRI modalities

MRSI



(c) Healthy



(d) CaP

Pros

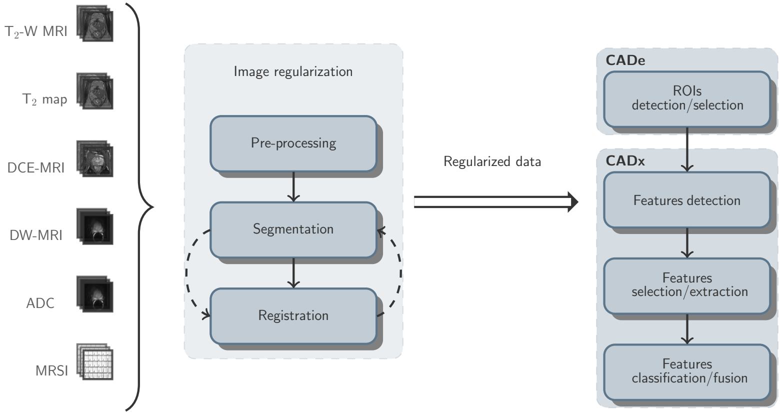
- Citrate correlated with Gleason score

Cons

- Low spatial resolution
- Variation inter-patients

CAD for CaP

Full CAD for detection and diagnosis of CaP

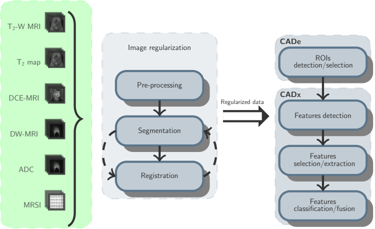


Common CAD framework based on MRI images used to detect CaP

CAD for CaP

Conclusions

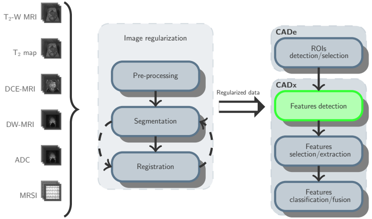
- ✓ 3 modalities better than 2
- ✓ Texture and edge features are predominant
- ✓ Features selection/extraction tends to improve performance
- ✓ Pre-eminence of SVM and ensemble classifier (i.e., AdaBoost, RF, etc.)



CAD for CaP

Conclusions

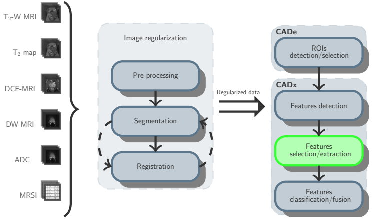
- ✓ 3 modalities better than 2
- ✓ Texture and edge features are predominant
- ✓ Features selection/extraction tends to improve performance
- ✓ Pre-eminence of SVM and ensemble classifier (i.e., AdaBoost, RF, etc.)



CAD for CaP

Conclusions

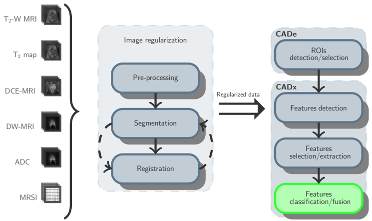
- ✓ 3 modalities better than 2
- ✓ Texture and edge features are predominant
- ✓ Features selection/extraction tends to improve performance
- ✓ Pre-eminence of SVM and ensemble classifier (i.e., AdaBoost, RF, etc.)



CAD for CaP

Conclusions

- ✓ 3 modalities better than 2
- ✓ Texture and edge features are predominant
- ✓ Features selection/extraction tends to improve performance
- ✓ Pre-eminence of SVM and ensemble classifier (i.e., AdaBoost, RF, etc.)



CAD for CaP

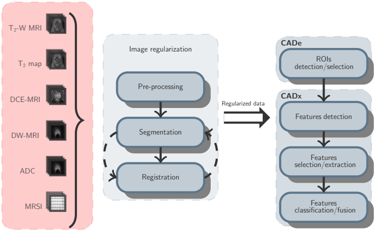
Conclusions

- ✓ 3 modalities better than 2
- ✓ Texture and edge features are predominant
- ✓ Features selection/extraction tends to improve performance
- ✓ Pre-eminence of SVM and ensemble classifier (i.e., AdaBoost, RF, etc.)

Scientific and technical challenges

- ✗ No publicly available mp-MRI dataset
- ✗ Only 1 study used 4 MRI modalities
- ✗ Limited work on data normalization
- ✗ A lot of features are extracted in 2D
- ✗ Limited work regarding selection/extraction
- ✗ No work regarding data balancing
- ✗ No source code available of any CAD

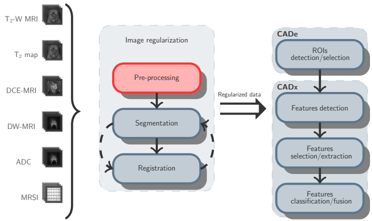
CAD for CaP



Scientific and technical challenges

- ✗ No publicly available mp-MRI dataset
- ✗ Only 1 study used 4 MRI modalities
- ✗ Limited work on data normalization
- ✗ A lot of features are extracted in 2D
- ✗ Limited work regarding selection/extraction
- ✗ No work regarding data balancing
- ✗ No source code available of any CAD

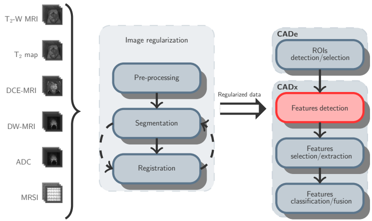
CAD for CaP



Scientific and technical challenges

- ✗ No publicly available mp-MRI dataset
- ✗ Only 1 study used 4 MRI modalities
- ✗ Limited work on data normalization
- ✗ A lot of features are extracted in 2D
- ✗ Limited work regarding selection/extraction
- ✗ No work regarding data balancing
- ✗ No source code available of any CAD

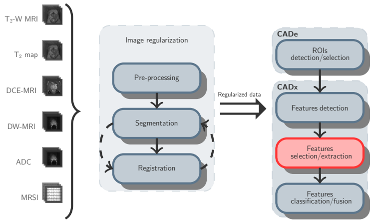
CAD for CaP



Scientific and technical challenges

- ✗ No publicly available mp-MRI dataset
- ✗ Only 1 study used 4 MRI modalities
- ✗ Limited work on data normalization
- ✗ A lot of features are extracted in 2D
- ✗ Limited work regarding selection/extraction
- ✗ No work regarding data balancing
- ✗ No source code available of any CAD

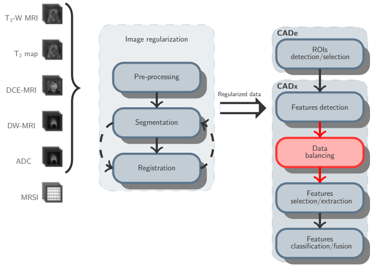
CAD for CaP



Scientific and technical challenges

- ✗ No publicly available mp-MRI dataset
- ✗ Only 1 study used 4 MRI modalities
- ✗ Limited work on data normalization
- ✗ A lot of features are extracted in 2D
- ✗ Limited work regarding selection/extraction
- ✗ No work regarding data balancing
- ✗ No source code available of any CAD

CAD for CaP



Scientific and technical challenges

- ✗ No publicly available mp-MRI dataset
- ✗ Only 1 study used 4 MRI modalities
- ✗ Limited work on data normalization
- ✗ A lot of features are extracted in 2D
- ✗ Limited work regarding selection/extraction
- ✗ No work regarding data balancing
- ✗ No source code available of any CAD

CAD for CaP

Conclusions

- ✓ 3 modalities better than 2
- ✓ Texture and edge features are predominant
- ✓ Features selection/extraction tends to improve performance
- ✓ Pre-eminence of SVM and ensemble classifier (i.e., AdaBoost, RF, etc.)

Scientific and technical challenges

- ✗ No publicly available mp-MRI dataset
- ✗ Only 1 study used 4 MRI modalities
- ✗ Limited work on data normalization
- ✗ A lot of features are extracted in 2D
- ✗ Limited work regarding selection/extraction
- ✗ No work regarding data balancing
- ✗ No source code available of any CAD

CAD for CaP

Conclusions

- ✓ 3 modalities better than 2
- ✓ Texture and edge features are predominant
- ✓ Features selection/extraction tends to improve performance
- ✓ Pre-eminence of SVM and ensemble classifier (i.e., AdaBoost, RF, etc.)

Scientific and technical challenges

- ✗ No publicly available mp-MRI dataset
- ✗ Only 1 study used 4 MRI modalities
- ✗ Limited work on data normalization
- ✗ A lot of features are extracted in 2D
- ✗ Limited work regarding selection/extraction
- ✗ No work regarding data balancing
- ✗ No source code available of any CAD

Research objectives

- Collect a mp-MRI dataset
- Design a CAD for CaP using all mp-MRI modalities
- Investigate normalization, feature selection/extraction, data balancing
- Implement 3D features
- Release source code and dataset

1 Introduction

2 State-of-the-art

3 I2CVB

- Mp-MRI prostate datasets
- Open source initiative
- I2CVB

4 DCE normalization

5 Toward a mp-MRI CAD for CaP

6 Experiments

7 Conclusions

Mp-MRI prostate datasets

1.5 T General Electric scanner

- T_2W-MRI, DW-MRI, DCE-MRI, and MRSI
- Ground-truth (GT) for CaP, PZ, and CG associated to T_2W-MRI modality
- Healthy: 4 *vs.* CaP: $\{$ PZ: $14+3$, CG: $0+3$ $\}$

3 T Siemens scanner

- T_2W-MRI, ADC, DCE-MRI, and MRSI
- GT for CaP, PZ, and CG associated to T_2W-MRI modality
- Additional GT of the prostate for DCE-MRI and ADC
- Healthy: 2 *vs.* CaP: $\{$ PZ: $12+2$, CG: $3+2$ $\}$

UdG

Mp-MRI prostate datasets

UBFC

1.5 T General Electric scanner

- T_2W-MRI, DW-MRI, DCE-MRI, and MRSI
- Ground-truth (GT) for CaP, PZ, and CG associated to T_2W-MRI modality
- Healthy: 4 *vs.* CaP: { PZ: $14+3$, CG: $0+3$ }

3 T Siemens scanner

- T_2W-MRI, ADC, DCE-MRI, and MRSI
- GT for CaP, PZ, and CG associated to T_2W-MRI modality
- Additional GT of the prostate for DCE-MRI and ADC
- Healthy: 2 *vs.* CaP: { PZ: $12+2$, CG: $3+2$ }

Open source initiative

protoclass toolbox

- Data management
- Features detection

imbalanced-learn toolbox[7]

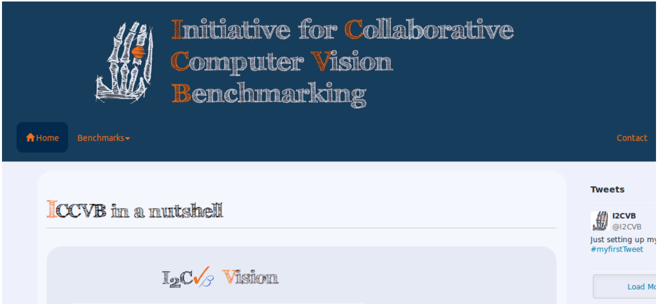
- Part of the `scikit-learn-contrib` projects

Third-party toolboxes

[7] Guillaume Lemaitre et al. "Imbalanced-learn: A Python Toolbox to Tackle the Curse of Imbalanced Datasets in Machine Learning". In: *Journal of Machine Learning Research* (2017).

A web platform

I_2CV_B platform

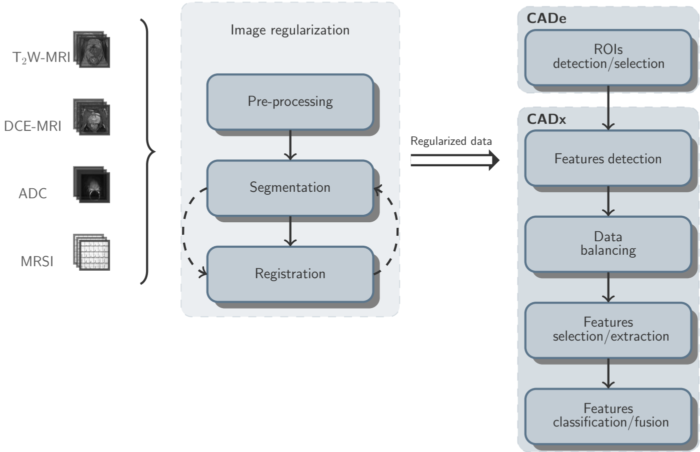


Hub for our different resources

- GitHub for our source codes
- Zenodo for our datasets
- HAL, arXiv, ResearchGate for our publications

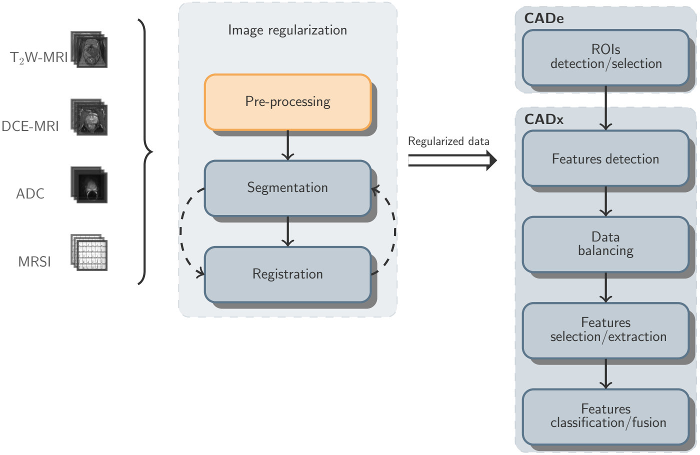
DCE-MRI normalization

Mp-MRI CAD for CaP



DCE-MRI normalization

Pre-processing

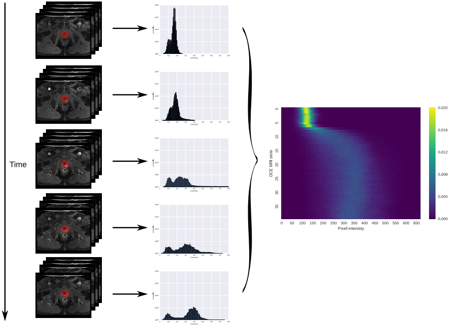


DCE-MRI normalization

Contribution[8]

- Propose a method to normalize DCE-MRI data

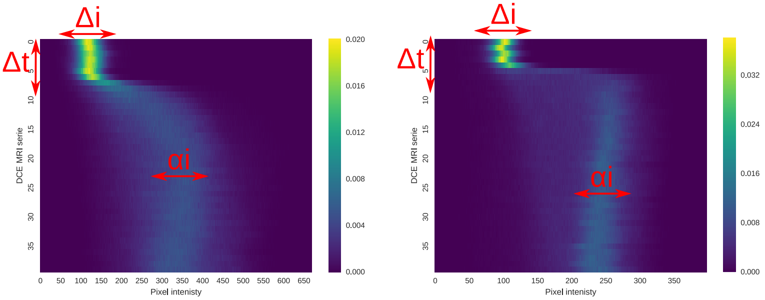
Heatmap representation



[8] Guillaume Lemaitre et al. "Automatic prostate cancer detection through DCE-MRI images: all you need is a good normalization". In: *Medical Image Analysis - Submitted* (2017).

DCE-MRI normalization

Inter-patients variations



(a) Patient #1 (b) Patient #2

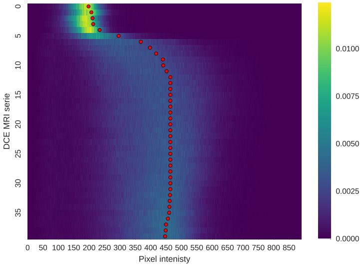
Variations driven by Δ_i, Δ_t, and σ_i

DCE-MRI normalization

Correction of Δ_i

Find the shortest path in a directed weighted graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ with the edge weight w_{ij} between 2 nodes i and j corresponding to 2 pixels at position (x_i, y_i) and (x_j, y_j), respectively defined as:

$$w_{ij} = \begin{cases} \alpha \exp(1 - \frac{H(i)}{\max(H)}) & \text{if } x_j = x_i + 1 \text{ and } y_j = y_i , \\ (1-\alpha) \exp(1 - \frac{H(i)}{\max(H)}) & \text{if } x_j = x_i \text{ and } y_j = y_i + 1 , \\ 0 & \text{otherwise,} \end{cases} \quad (1)$$



DCE-MRI normalization

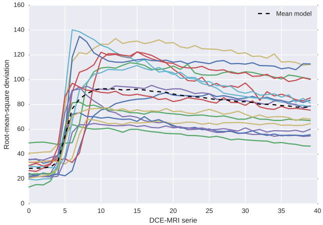
Correction of Δ_t and σ_i

Register all RMSD to a mean model such that:

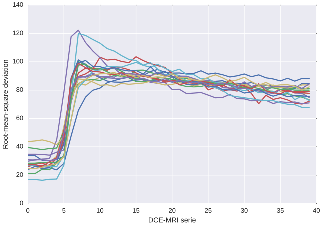
$$\underset{\alpha,\tau}{\arg\min} = \sum_{t=1}^{N} \left[T\left(\alpha, \tau, f(t)\right) - \mu(t) \right]^2 , \tag{2}$$

$$f(t) = \sqrt{\left(\frac{\sum_{n=1}^{N} x(t)_n^2}{N} \right)}, \tag{3}$$

$$T(\alpha, \tau, f(t)) = \alpha f(t - \tau). \tag{4}$$



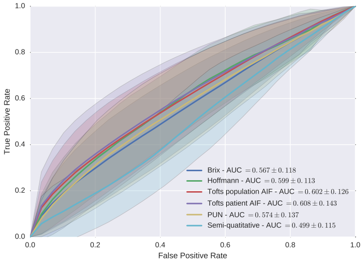
(a) RMSD before correction



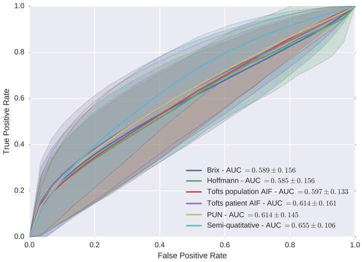
(b) Registered RMSD

DCE modality

Quantitative and semi-quantitative models



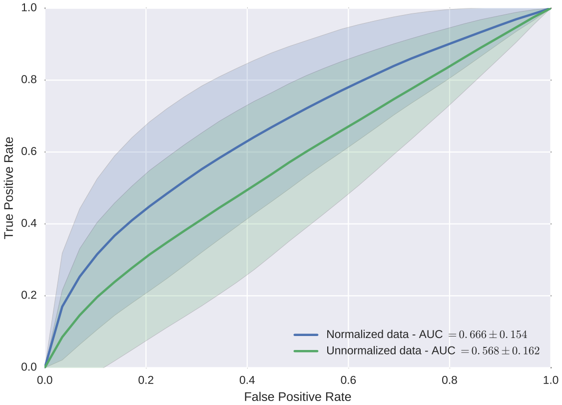
(a) Without normalization



(b) With normalization

DCE modality

Entire signal



DCE Normalization

Mp-MRI CAD for CaP

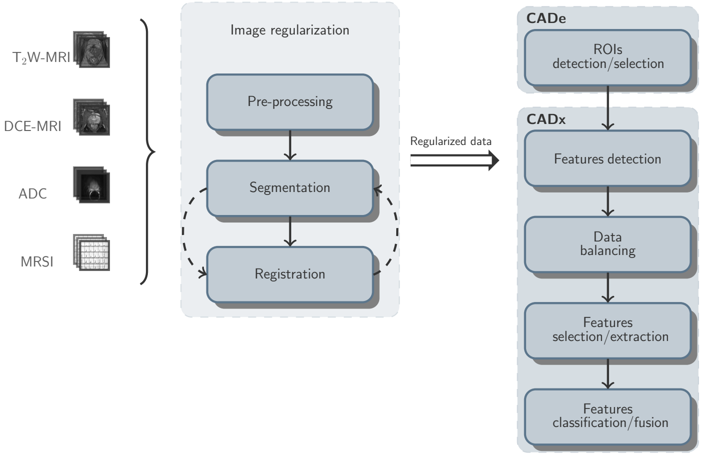
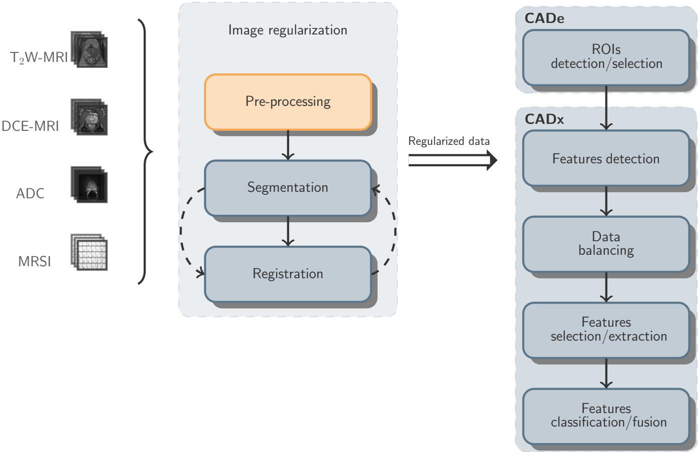


Image regularization

Pre-processing



Pre-processing

T_2W-MRI normalization

- Rician normalization[9]

ADC normalization

- Piecewise-linear normalization

DCE-MRI normalization

- Graph and deviation based normalization[10]

MRSI normalization

- Phase correction[11]
- Frequency alignment
- Baseline correction[12]

[10] Guillaume Lemaitre et al. "Normalization of T2W-MRI Prostate Images using Rician a priori". In: *SPIE Medical Imaging*. International Society for Optics and Photonics. 2016, pp. 978529–978529.

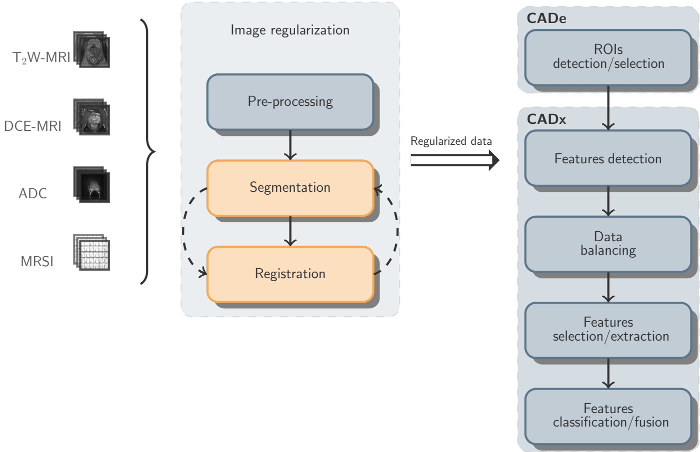
[11] Lemaitre et al., "Automatic prostate cancer detection through DCE-MRI images: all you need is a good normalization".

[12] Li Chen et al. "An efficient algorithm for automatic phase correction of {NMR} spectra based on entropy minimization ". In: *Journal of Magnetic Resonance* 158.12 (2002), pp. 164–168.

[13] Yuanxin Xi and David M Rocke. "Baseline correction for NMR spectroscopic metabolomics data analysis". In: *BMC bioinformatics* 9.1 (2008), p. 1.

Image regularization

Segmentation & registration



UdG

Segmentation & registration

UBFC

Interpolation

- ADC and DCE-MRI are interpolated to the T_2W-MRI resolution

Segmentation

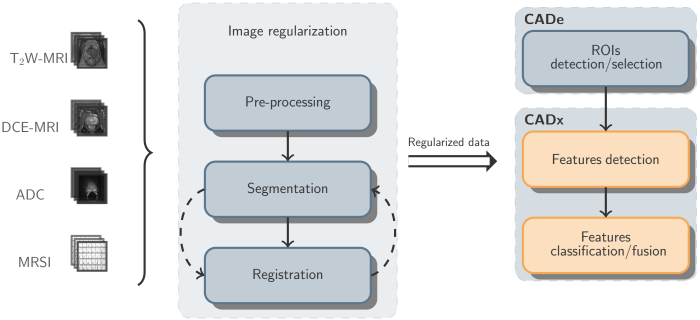
- Manual prostate segmentation available for T_2W-MRI, DCE-MRI, and ADC
- Cap, PZ, and CG manual segmentation available for T_2W-MRI

Registration

- Intra-patient motions correction in DCE-MRI: rigid registration using mutual information
- DCE-MRI is registered to T_2W-MRI using the prostate segmentation
- ADC is registered to T_2W-MRI using the prostate segmentation

CADe-CADx

Features detection



Feature detection

T_2W-MRI and ADC features

- Intensity
- Kirsch filter
- Laplacian filter
- Prewitt filter
- Scharr filter
- Sobel filter
- DCT decomposition
- Gabor filters
- Phase congruency filter
- Haralick filter
- LBP filter

DCE-MRI features

- Brix's and Hoffmann's model
- Tofts' model
- PUN model
- Semi-quantitative model
- Entire enhanced signal

MRSI features

- Quantification with fixed bounds
- Quantification by fitting some modeled signal
- Entire spectra

Anatomical features

Metrics[14],[15]

- Relative distance to the *prostate boundary*
- Relative distance to the *prostate center*
- Relative position in the *Euclidean* coordinate systems
- Relative position in the *cylindrical* coordinate systems

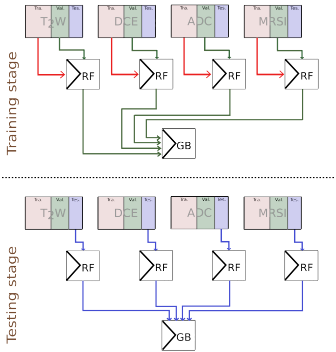
[14] Jeremy Chen et al. "Automatic determination of arterial input function for dynamic contrast enhanced MRI in tumor assessment". In: *International Conference on Medical Image Computing and Computer-Assisted Intervention*. Springer. 2008, pp. 594–601. DOI: 10.1007/978-3-540-85988-8_71.

[15] G. Litjens et al. "Computer-iided detection of prostate cancer in MRI". . In: *Medical Imaging, IEEE Transactions on* 33.5 (May 2014), pp. 1083–1092. ISSN: 0278-0062.

Features classification

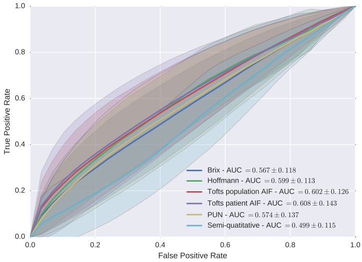
Classifier

- Random forest (RF)
- Stacking of RF with an adaboost and gradient-boosting meta-classifier

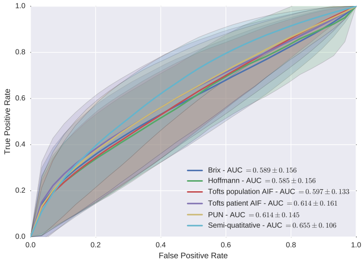


DCE modality

Quantitative and semi-quantitative models



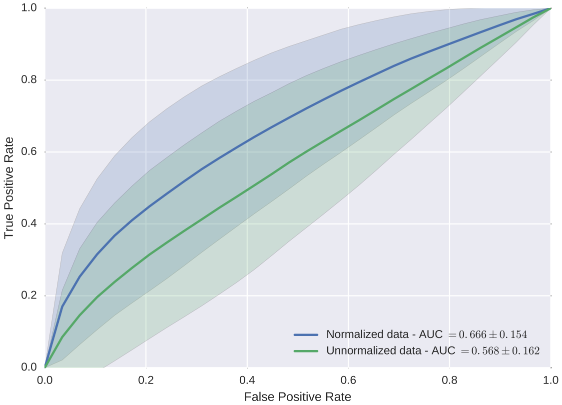
(a) Without normalization



(b) With normalization

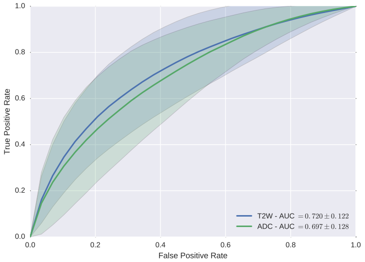
DCE modality

Entire signal

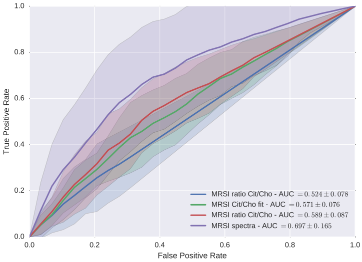


T_2W-MRI, ADC, and MRSI modalities

ROC analysis



(a) T_2W-MRI and ADC



(b) MRSI

T_2W-MRI, ADC, and MRSI modalities

Discussions

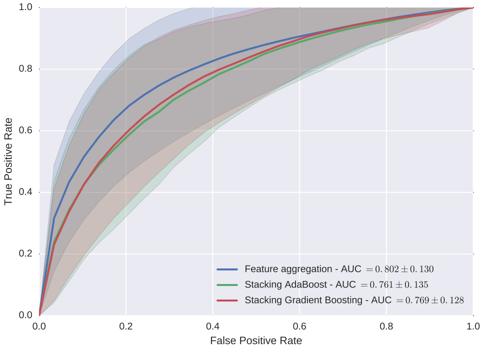
- DCE-MRI: normalized data $\rightarrow$ best performance
- DCE-MRI: entire signal better than models
- MRSI: fitting better than bounds approach
- MRSI: entire spectra better than others
- T_2W-MRI $<$ ADC $=$ MRSI $<$ DCE
- Performance at an "acceptable" level of discrimination - AUC $\in [0.7, 0.8]$

Conclusions

- ✓ DCE-MRI: Use the normalized enhanced signal
- ✓ MRSI: Use the entire spectra

Coarse combination

Aggregation *vs.* stacking

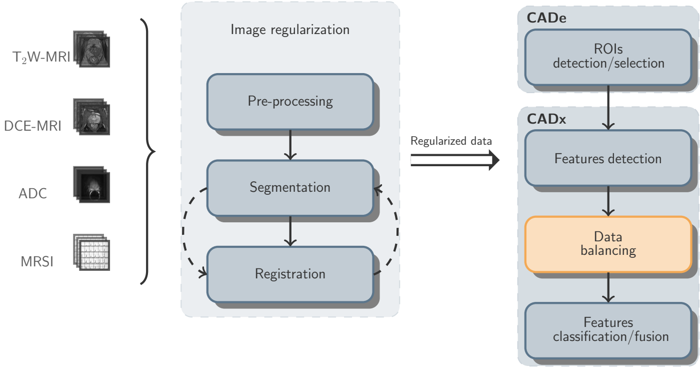


Conclusions

- ▶ Aggregation leads to best performance
- ✓ Gradient boosting is preferable to adaboost

CADe-CADx

Data balancing

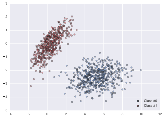


Data balancing

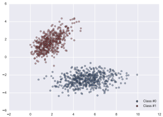
Toy example

Data balancing

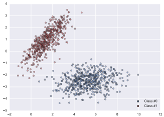
Under-sampling



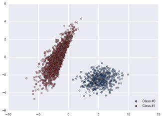
(a) NM1



(b) NM2

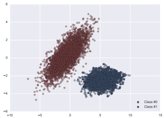


(c) NM3

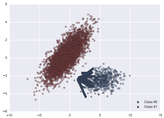


(d) IHT

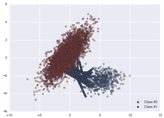
Over-sampling



(e) SMOTE



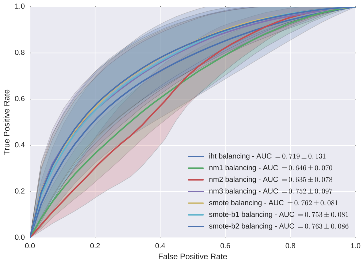
(f) SMOTE-b1



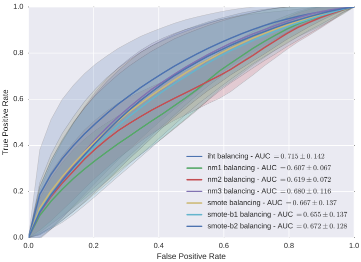
(g) SMOTE-b2

Data balancing

T_2W-MRI and ADC



(a) T_2W-MRI



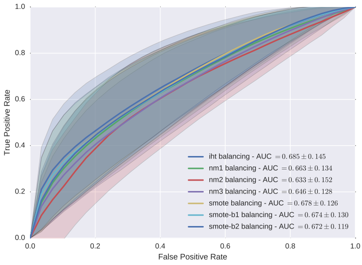
(b) ADC

Conclusions

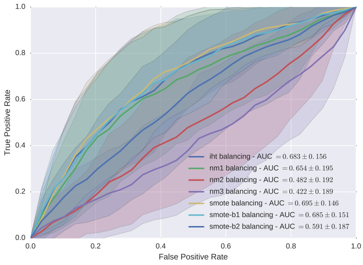
- ✓ IHT $\rightarrow$ ADC
- ✓ SMOTE $\rightarrow$ T_2W-MRI

Data balancing

DCE-MRI and MRSI



(c) DCE-MRI



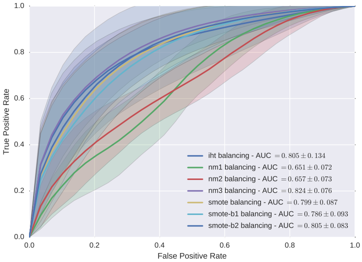
(d) MRSI

Conclusions

- ✓ IHT $\rightarrow$ ADC and DCE-MRI
- ✓ SMOTE $\rightarrow$ T_2W-MRI and MRSI

Data balancing

Aggregation

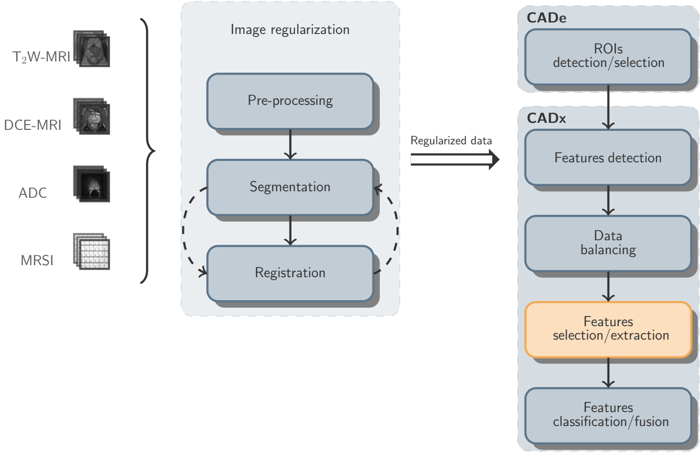


Conclusions

- ✓ IHT $\rightarrow$ ADC and DCE-MRI
- ✓ SMOTE $\rightarrow$ T_2W-MRI and MRSI
- ✓ NM3 $\rightarrow$ aggregate feature

CADe-CADx

Features selection/extraction



Features selection/extraction

Features extraction

- Independent components analysis (ICA)
- Principal components analysis (PCA)
- Sparse-PCA

Features selection

- One-way analysis of variance (ANOVA)
- Gini importance

UdG

Features selection/extraction

UBFC

T_2W-MRI

Methods	Percentiles						
	15	17.5	20	22.5	25	27.5	30
ANOVA F-score	0.755 ± 0.049	0.770 ± 0.058	0.777 ± 0.064	0.782 ± 0.066	$\mathbf{0.784} \pm \mathbf{0.067}$	0.783 ± 0.072	0.782 ± 0.070

Methods	Percentiles						
	1	2	5	10	15	20	30
Gini importance	0.726 ± 0.064	0.731 ± 0.055	0.751 ± 0.065	0.758 ± 0.076	0.752 ± 0.087	0.761 ± 0.077	$\mathbf{0.764} \pm \mathbf{0.079}$

Features selection/extraction

ADC

Methods	Percentiles						
	10	12.5	15	17.5	20	22.5	25
ANOVA F-score	0.684 ± 0.123	0.713 ± 0.125	0.712 ± 0.134	0.710 ± 0.144	$\mathbf{0.714} \pm \mathbf{0.142}$	0.708 ± 0.150	0.708 ± 0.150

Methods	Percentiles						
	1	2	5	10	15	20	30
Gini importance	0.672 ± 0.132	0.690 ± 0.138	$\mathbf{0.743} \pm \mathbf{0.139}$	0.730 ± 0.136	0.730 ± 0.142	0.724 ± 0.141	0.722 ± 0.142

UdG

Features selection/extraction

UBFC

DCE-MRI

Methods	Number of components or sparsity level						
	2	4	8	16	24	32	36
PCA	0.656 ± 0.133	0.634 ± 0.121	0.668 ± 0.149	0.680 ± 0.145	0.682 ± 0.146	0.679 ± 0.151	0.683 ± 0.149
Sparse-PCA	0.578 ± 0.117	0.546 ± 0.121	0.554 ± 0.097	—	—	—	—
ICA	0.657 ± 0.132	0.629 ± 0.117	0.671 ± 0.157	0.686 ± 0.158	$\mathbf{0.691} \pm \mathbf{0.158}$	0.681 ± 0.161	0.679 ± 0.166

Features selection/extraction

MRSI

Methods	Number of components or sparsity level						
	2	4	8	16	24	32	36
PCA	0.566 ± 0.120	0.575 ± 0.141	0.648 ± 0.162	0.662 ± 0.177	0.659 ± 0.184	0.671 ± 0.179	0.672 ± 0.182
Sparse-PCA	0.502 ± 0.050	0.571 ± 0.158	0.585 ± 0.111	—	—	—	—
ICA	0.567 ± 0.119	0.578 ± 0.140	0.654 ± 0.145	0.656 ± 0.167	0.650 ± 0.187	0.663 ± 0.174	**0.677 ± 0.171**

Features selection/extraction

Aggregation

Methods	**Percentiles**						
	10	12.5	15	17.5	20	22.5	25
ANOVA F-score	0.764 ± 0.095	0.765 ± 0.079	0.800 ± 0.083	0.817 ± 0.089	$\mathbf{0.828} \pm \mathbf{0.084}$	$0.822 \pm 0.0.084$	0.815 ± 0.086

Methods	**Percentiles**						
	10	12.5	15	17.5	20	22.5	25
Gini importance	0.834 ± 0.085	0.834 ± 0.088	0.834 ± 0.084	$\mathbf{0.836} \pm \mathbf{0.083}$	0.834 ± 0.079	0.828 ± 0.086	0.830 ± 0.077

Features selection/extraction

Conclusions

- ✓ T_2W-MRI: ANOVA-based selection with 25 % of the data
- ✓ ADC: Gini importance-based selection with 5 % of the data
- ✓ DCE-MRI: ICA with 24 components
- ✓ MRSI: ICA with 36 components
- ✓ Aggregation: Gini importance with 17.5 % of the data

Features selection/extraction

Selected features in T_2W-MRI and ADC

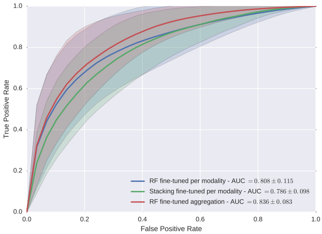
T_2W-MRI	ADC
8 edges 155 Gabor filters 2 Haralick features 1 intensity 4 LBP 2 phase congruency	1 DCT 32 Gabor filters 1 phase congruency
172 features	**34 features**

Selected features with aggregation

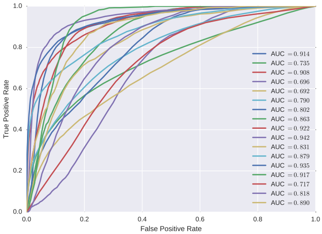
T_2W-MRI	ADC	DCE-MRI	MRSI
113 Gabor filters 1 phase congruency 4 edges 1 intensity	53 Gabor filters 2 phase congruency	14 samples	78 samples
267 features			

Fine-tuned combination

Aggregation *vs.* stacking



ROC for each patient

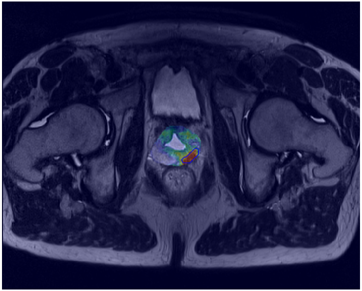


UdG

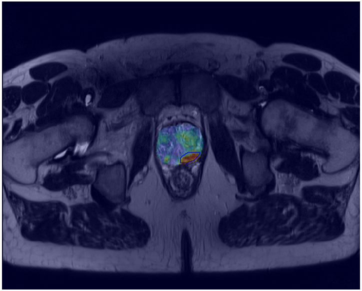
Fine-tuned combination

UBFC

"Outstanding" discrimination level



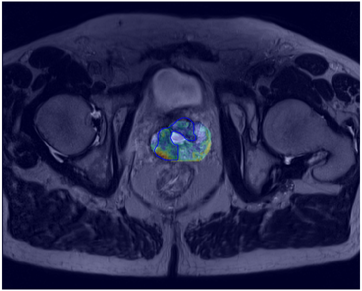
(e) AUC $= 0.922$



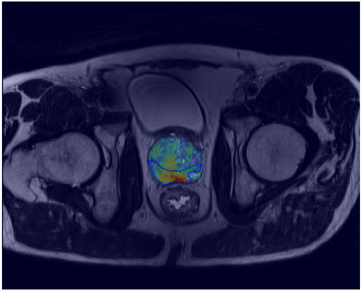
(f) AUC $= 0.914$

Fine-tuned combination

"Acceptable" discrimination level



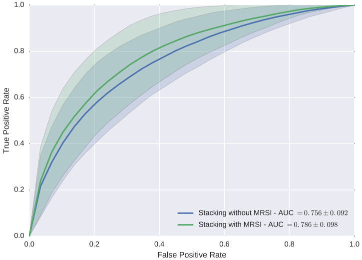
(g) AUC $= 0.692$



(h) AUC $= 0.735$

MRSI benefit

Stacking with/without MRSI



MRSI in aggregation

✓ Features from MRSI are the most selected features

Contributions & future works

Research objectives

- Collect a mp-MRI dataset
- Design a CAD for CaP using all mp-MRI modalities
- Investigate normalization, feature selection/extraction, data balancing
- Implement 3D features
- Release source code and dataset

Contributions & future works

Contributions

- ✓ Collect a mp-MRI dataset
- ✓ Design a CAD for CaP using all mp-MRI modalities
- ✓ Investigate normalization, feature selection/extraction, data balancing
- ✓ Implement 3D features
- ✓ Release source code and dataset

Contributions & future works

Avenue for future research

- ✗ Incorporate spatial information in classification using super-voxels
- ✗ Dissociate classifiers for the PZ and CG regions
- ✗ Explore the *radiomics* features from PI-RADS v.2
- ✗ Investigate the benefit of deep-learning

Timeline