Computer-Aided Diagnosis for Prostate Cancer using mp-MRI

PhD Defence
28th November 2016

Guillaume Lemaître

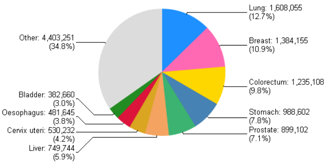
Universitat de Girona - ViCOROB
Université de Bourgogne Franche-Comté - LE2I

Supervised by:

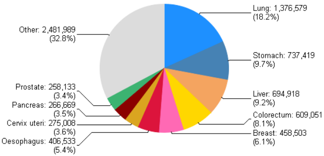
Robert Martí - Fabrice Mériaudeau
Jordi Freixenet - Paul M. Walker

Motivations

Statistics



(a) # of cancer cases



(b) # of cancer deaths

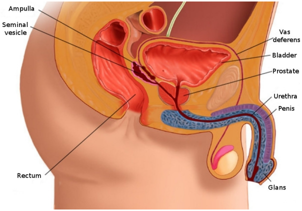
Implications, image source[1]

- 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]

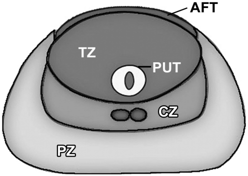
Characteristics

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

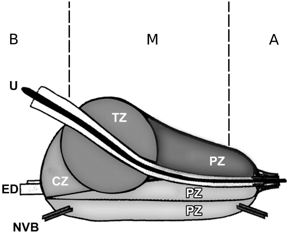
[2] Geckomedia. *Natom Anatomy*. French. June 2011. url: http://www.natomshop.com/.

The prostate organ

Anatomy



(a) Transverse plane

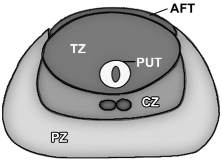


(b) Sagittal plane

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 → 85 %
- Fast-growing → 15 %
- CaPs in CG are more aggressive

Zonal predisposition

- PZ → 70 % to 80 %
- TZ → 10 % to 20 %
- CG → 5 %

Goals

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

Screening

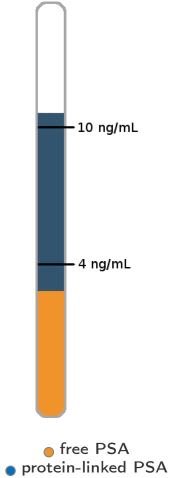
Prostate-specific antigen

- $> 10\,\mathrm{ng\,mL^{-1}} \rightarrow$ biopsy
- From $4\,\mathrm{ng\,mL^{-1}}$ to $10\,\mathrm{ng\,mL^{-1}}$

  $\rightarrow \frac{\bullet}{\bullet + \bullet} > 15\% \rightarrow$ biopsy
- ✗ Not reliable

"Blind" transrectal ultrasound biopsy

- Take samples from different locations
- Grade using Gleason score
- ✗ Invasive procedure
- ✗ Lead to false positives & negatives



Screening

Prostate-specific antigen

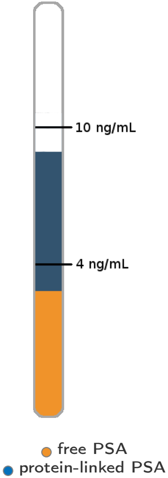
- $> 10\,\text{ng}\,\text{mL}^{-1} \rightarrow$ biopsy
- From $4\,\text{ng}\,\text{mL}^{-1}$ to $10\,\text{ng}\,\text{mL}^{-1}$
  $\rightarrow \frac{\bullet}{\bullet + \bullet} > 15\% \rightarrow$ biopsy

✗ Not reliable

"Blind" transrectal ultrasound biopsy

- Take samples from different locations
- Grade using Gleason score

✗ Invasive procedure
✗ Lead to false positives & negatives



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

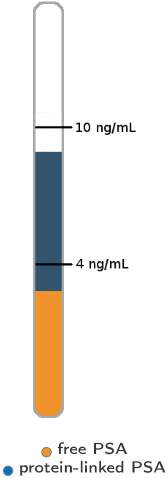
✗ Not reliable

"Blind" transrectal ultrasound biopsy

- Take samples from different locations
- Grade using Gleason score

✗ Invasive procedure

✗ Lead to false positives & negatives



UdG

Screening

UBFC

Prostate-specific antigen

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

✗ Not reliable

"Blind" transrectal ultrasound biopsy

- Take samples from different locations
- Grade using Gleason score

✗ Invasive procedure

✗ Lead to false positives & negatives

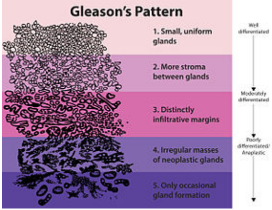


Image source: `https://goo.gl/fEVQXQ`

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
- ✗ Not reliable

"Blind" transrectal ultrasound biopsy

- Take samples from different locations
- Grade using Gleason score
- ✗ Invasive procedure
- ✗ Lead to false positives & negatives

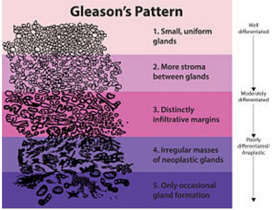


Image source: `https://goo.gl/fEVQXQ`

UdG

Screening

UBFC

Prostate-specific antigen

- $> 10\,\mathrm{ng\,mL^{-1}} \rightarrow$ biopsy
- From $4\,\mathrm{ng\,mL^{-1}}$ to $10\,\mathrm{ng\,mL^{-1}}$

  $\rightarrow \frac{\bullet}{\bullet + \bullet} > 15\% \rightarrow$ biopsy

✗ Not reliable

“Blind” transrectal ultrasound biopsy

- Take samples from different locations
- Grade using Gleason score

✗ Invasive procedure

✗ Lead to false positives & negatives

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. Schröder 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

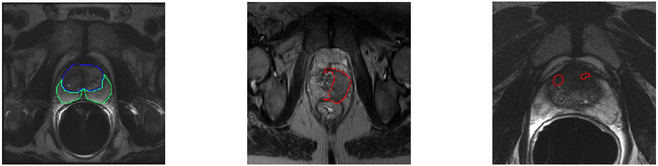
Propose a mp-MRI CAD for CaP

- 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

UdG · UBFC

MRI modalities

T_2W-MRI



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

Healthy

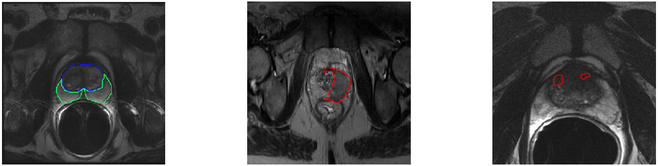
- Intermediate to high-signal intensity (SI) in PZ
- Low-SI in CG

CaP

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

MRI modalities

T_2W-MRI



(d) Healthy (e) CaP PZ (f) CaP CG

Pros

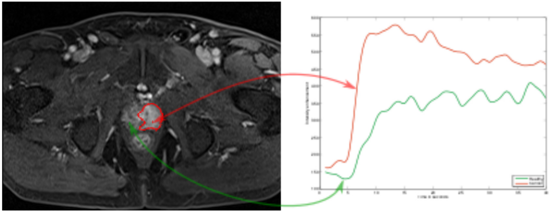
- Highest spatial resolution
- Anatomy nicely depicted

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

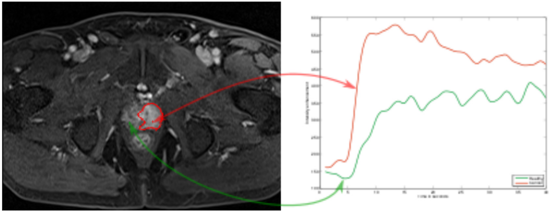
- Faster wash-in, wash-out, time-to-peak enhancement
- Higher integral under the curve, max SI

UdG

MRI modalities

UBFC

DCE-MRI



Green: healthy - Red: CaP

Pros

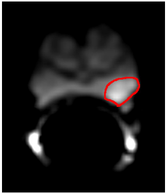
- Information about vascularity

Cons

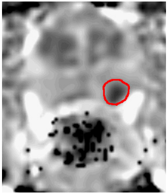
- Spatial mis-registration
- Lower spatial resolution than T_2W-MRI
- Difficult detection in CG

MRI modalities

DW-MRI - ADC



(a) DW MRI



(b) ADC

Healthy

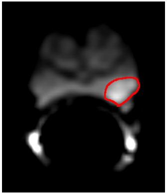
- DW-MRI: lower SI
- ADC: higher-SI

CaP

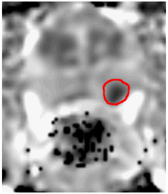
- DW-MRI: higher SI
- ADC: lower-SI

MRI modalities

DW-MRI - ADC



(c) DW MRI



(d) ADC

Pros

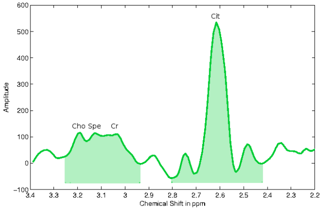
- Information about tissue structure
- ADC correlated with Gleason score

Cons

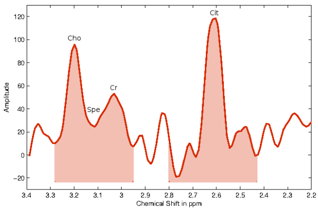
- Poor spatial resolution
- Variability of the ADC coefficient

MRI modalities

MRSI



(a) Healthy



(b) CaP

Healthy

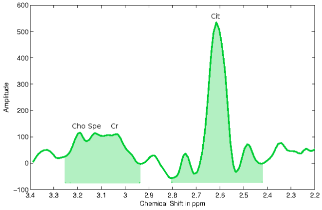
- High citrate
- Moderate choline and spermine

CaP

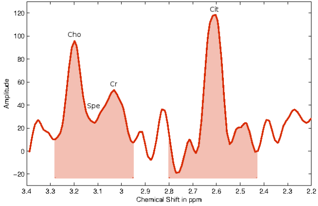
- Decrease of citrate and spermine
- Increase of choline

MRI modalities

MRSI



(c) Healthy



(d) CaP

Pros

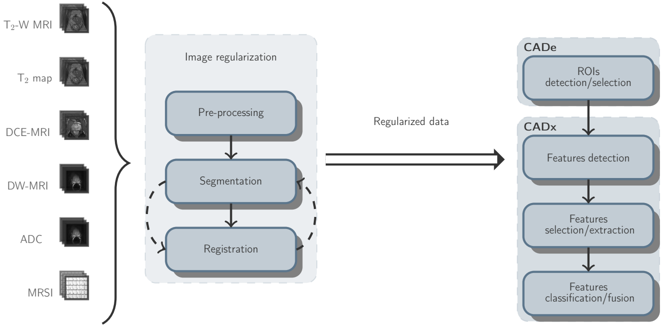
- Citrate correlated with Gleason score

Cons

- Low spatial resolution
- Variation inter-patients

CAD for CaP

Full CAD for detection and diagnosis of CaP

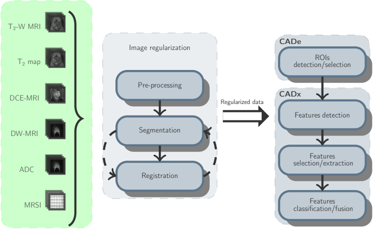


Common CAD framework based on MRI images used to detect CaP

CAD for CaP

Conclusions

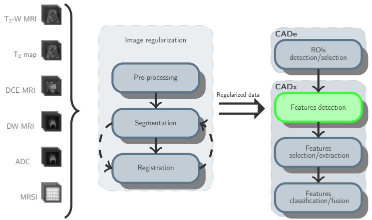
- ✓ 3 modalities better than 2
- ✓ Texture and edge features are predominant
- ✓ Features selection/extraction tends to improve performance
- ✓ Pre-eminence of SVM and ensemble classifier (i.e., AdaBoost, RF, etc.)



CAD for CaP

Conclusions

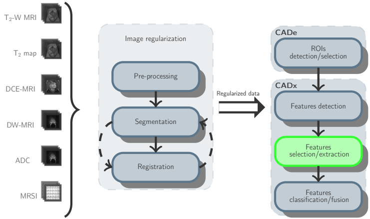
- ✓ 3 modalities better than 2
- ✓ Texture and edge features are predominant
- ✓ Features selection/extraction tends to improve performance
- ✓ Pre-eminence of SVM and ensemble classifier (i.e., AdaBoost, RF, etc.)



CAD for CaP

Conclusions

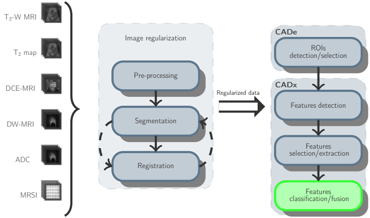
- ✓ 3 modalities better than 2
- ✓ Texture and edge features are predominant
- ✓ Features selection/extraction tends to improve performance
- ✓ Pre-eminence of SVM and ensemble classifier (i.e., AdaBoost, RF, etc.)



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

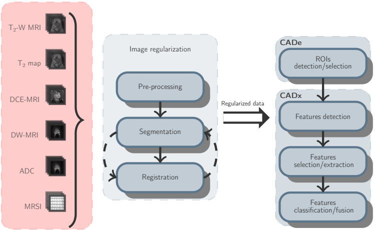
Conclusions

- ✓ 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

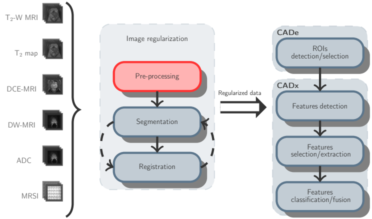
CAD for CaP



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

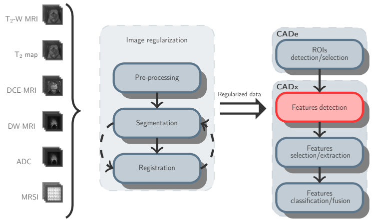
CAD for CaP



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

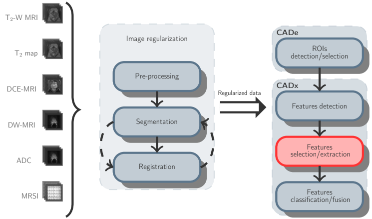
CAD for CaP



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

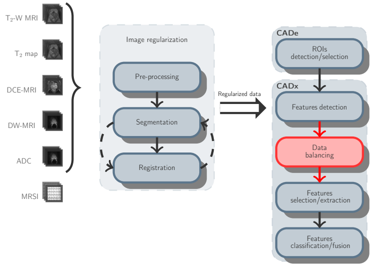
CAD for CaP



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

UdG

CAD for CaP

UBFC

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

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.)

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

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: $3 + 2$ $\}$

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

Open source initiative

UBFC

protoclass toolbox

- Data management
- Features detection

imbalanced-learn toolbox[7]

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

Third-party toolboxes

NumPy SciPy matplotlib

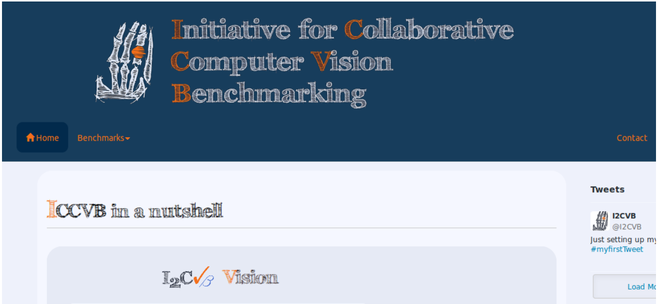
IP[y]: IPython Interactive Computing

SimpleITK scikit-image image processing in python scikit learn

[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

I2CVB platform

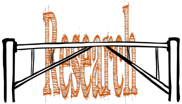


Hub for our different resources

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

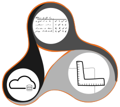
Manifesto

Vision



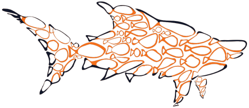
- Ease the access to make research

Mission



- Open data; evaluation methods; comparison framework; reporting platform

Protagonists



- Research groups and individuals from all walks of life to shape an open community

Strategy

- Use successful practises from Free Software and Quality Management

1. Introduction
2. State-of-the-art
3. I2CVB
4. Toward a mp-MRI CAD for CaP
   - Image regularization
     - T_2W-MRI normalization
     - DCE-MRI normalization
     - MRSI pre-processing
     - Segmentation & registration
   - CADe-CADx
     - Features detection
     - Data balancing
     - Features selection/extraction
     - Features classification
5. Experiments & validation

Toward a mp-MRI CAD for CaP

Mp-MRI CAD for CaP

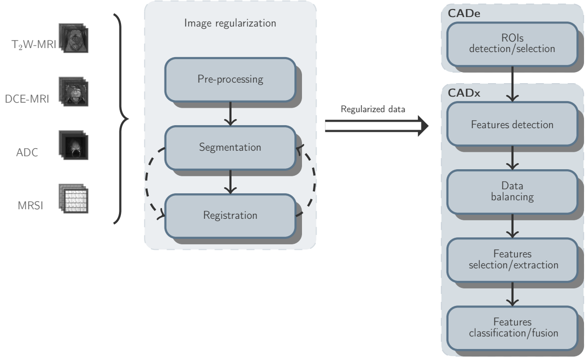
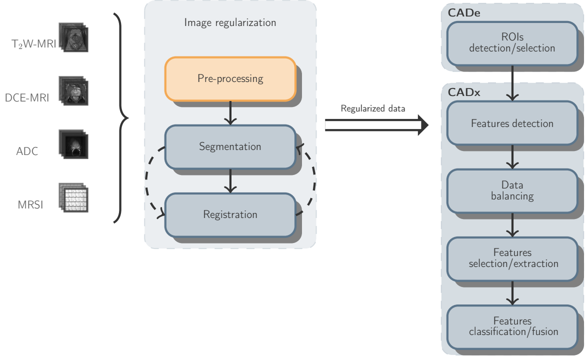


Image regularization

Pre-processing



T_2W-MRI normalization

UdG

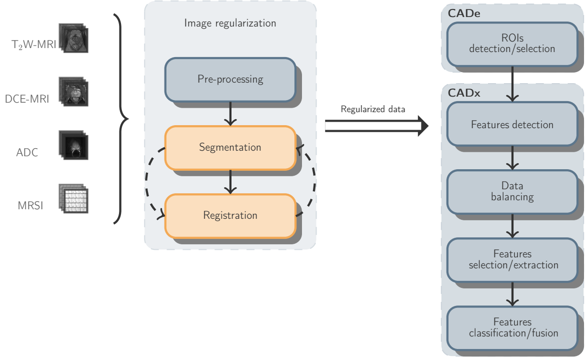
DCE-MRI normalization

UBFC

MRSI pre-processing

Image regularization

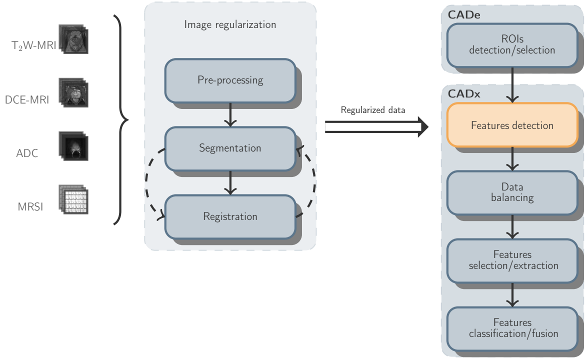
Segmentation & registration



Segmentation & registration

CADe-CADx

Features detection



T_2W-MRI and ADC map

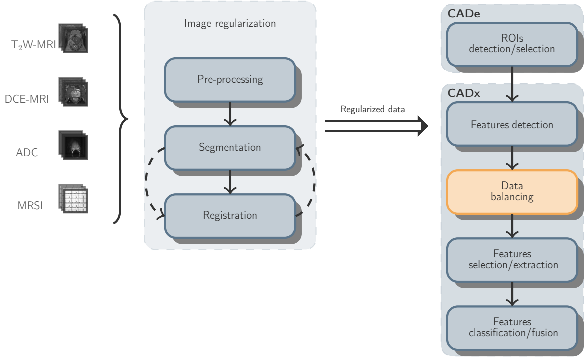
DCE-MRI

MRSI

Anatomical features

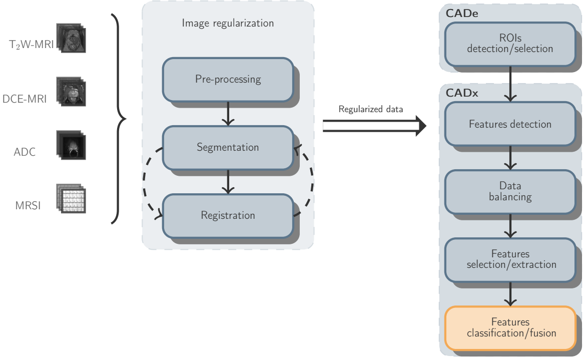
CADe-CADx

Data balancing



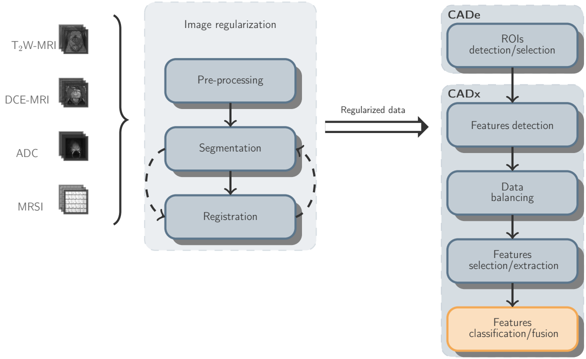
CADe-CADx

Features selection/extraction



CADe-CADx

Features classification



T_2W-MRI normalization

DCE-MRI normalization

Standalone modalities

Coarse combination

Data balancing

Features selection/extraction

UdG

Fine-tuned combination

UBFC

UdG

MRSI benefit

UBFC